

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 16:56:01 ; Search time 1223 Seconds  
(without alignments)  
11515.101 Million cell updates/sec

Title: US-09-996-956-1

Perfect score: 5217

Sequence: 1 gaacttttaaatatccctc.....taaaaaaaaaaaaaaaaaaa 5217

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2525.4	48.4	11801	21 AAC61388	An androgen-related
2	1531.6	29.4	1579	21 AAC61382	cDNA sequence of a
3	1456.8	27.9	1603	21 AAC61381	cDNA sequence of a
C 4	1308	25.1	1804	23 ABV23954	Human prostate exp
C 5	1308	25.1	1804	23 ABV24032	Human prostate exp
C 6	1308	25.1	1804	23 ABV29837	Human prostate exp
C 7	1308	25.1	1804	23 ABV29916	Human prostate exp
8	891.4	17.1	1025	21 AAC61384	Probe for androgen

```
c 9 462 8.9 249999 25 ABZ80229 Human tramdorin ge
10 454 8.7 2612 22 AAH17215 Human cDNA sequenc
11 438.4 8.4 843 24 ABQ89916 Human prostate exp
c 12 421.2 8.1 33206 24 ABN95882 Gene #2380 used to
13 417.6 8.0 1761 21 AAC74309 Human secreted pro
c 14 404.6 7.8 80240 20 AAV83940 NC-contlig derived
c 15 404.6 7.8 80595 20 AAV83939 HC-contlig derived
c 16 390.6 7.5 27666 25 ABZ74190 Secreted protein g
17 390.6 7.5 27666 25 ABZ67756 Human secreted pro
c 18 390.6 7.5 32212 22 AAL06082 Human reproductive
c 19 390.6 7.5 32212 23 ABL98647 Human testicular a
c 20 390.6 7.5 41206 25 ABZ74191 Secreted protein g
21 390.6 7.5 41206 25 ABZ67757 Human secreted pro
22 390.4 7.5 848 24 ABQ89711 Human prostate exp
23 386.6 7.4 820 24 ABQ89731 Human prostate exp
c 24 386.2 7.4 6138 22 AAL05378 Human reproductive
c 25 386.2 7.4 6138 23 ABL98244 Human testicular a
26 380.4 7.3 790 24 ABQ89552 Human prostate exp
27 377.8 7.2 802 24 ABQ89827 Human prostate exp
28 374.8 7.2 237961 24 ABQ80552 Human Carlon gene
29 363.6 7.0 1691080 24 ABX08336 Human phospholiest
30 359.8 6.9 599 22 ABQ62069 Human foetal liver
31 359.8 6.9 599 22 ABA29520 Probe #7986 for ge
32 359.8 6.9 599 22 AAK10387 Human brain expres
33 359.8 6.9 599 22 AAK86288 Human bone marrow
34 359.8 6.9 599 22 AAI17277 Probe #7210 for ge
35 359.8 6.9 599 22 AAI42011 Probe #10697 used
36 359.8 6.9 599 23 ABS35968 Human liver single
37 359.8 6.9 599 24 ABS10371 Human genome-deriv
c 38 357.8 6.9 438 23 ABV36282 Human prostate exp
c 39 357.8 6.9 438 23 ABV45318 Human prostate exp
c 40 357.4 6.9 409 23 ABV15490 Human prostate exp
c 41 351 6.7 1650 24 ABL49826 Human Ras GTP enzy
c 42 350.8 6.7 1980 22 AAK94596 Human full-length
c 43 350 6.7 28818 22 AAL35901 Human musculoskele
44 350 6.7 28818 25 ABX58889 cDNA encoding nove
45 338.8 6.5 154465 24 AAD28763 Human AKAP allelic
```

#### ALIGNMENTS

```
RESULT 1
AAC61388
ID AAC61388 standard; DNA; 11801 BP.
AC AAC61388;
XX
XX 19-FEB-2001 (first entry)
DT
DE An androgen-related, prostate-specific gene PCGEM1.
XX
XX Androgen-regulated gene; prostate specific gene; PCGEM1; prostate cancer;
KW prostate cancer gene expression marker 1; prostate related disease;
KW benign prostatic hyperplasia; ss.
XX
XX Homo sapiens.
XX
XX WO200058470-A1.
XX
XX 05-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US07906.
XX
XX 26-MAR-1999; 99US-0126469.
XX
XX (SRK/) SRIKANTAN V.
PA (ZOUZ/) ZOU Z.
PA (MOUL/) MOUL J W.
PA (SRIV/) SRIVASTAVA S.
XX
XX Srikantan V, Zou Z, Moul JW, Srivastava S;
```



Db 10721 TAGATTTTGTATTTTCCACATGTTGGCCAGGCTGGTCTCGAAGCCCTAGCTCAGGTGA 10780  
Qy 2211 TCCCTCGCTCGGCTCTCAAGTGTGGGATACAGGTGTGAGCCACCGCCAGCC 2270  
Db 10781 TCCCTCGCTCGGCTCTCAAGTGTGGGATACAGGTGTGAGCCACCGCCAGCC 10840  
Qy 2271 AATATCAGTTTCTTTTAAACACAAAGCTAAACAAATCAAAATACCTAGTAGGGGAGA 2330  
Db 10841 AATATCAGTTTCTTTTAAACACAAAGCTAAACAAATCAAAATACCTAGTAGGGGAGA 10900  
Qy 2331 AAAAAAATAAGGCACTGTTTATGTGTAAACAGGCTCTGTGTCAATCAGTGGG---CA 2386  
Db 10901 AAAAAAATAAGGCACTGTTTATGTGTAAACAGGCTCTGTGTCAATCAGTGGGCGAGA 10960  
Qy 2387 GACATAAACAGTAAAGATC-AATCCTTTTTCATATATCC---TTCCTGCAGAAATACATAA 2442  
Db 10961 CCAATAAACAGTAAAGATCAAAATCCTTTTCATATAATCCTTTTCGCAAAATACATAA 11020  
Qy 2443 AATCCC--ACAAATGGCTATCTTCTCTTTTATGATAT-TTGGAGAAATGTAGCTAAAGTGA 2499  
Db 11021 AATCCCAACAATGGCTTATCTTCTCTTTTATGATATGTGTGAGAAATGTAGCTAAAGTGA 11080  
Qy 2500 CAGATATTTTCTTGGGTGTATAGACCACAAAGACGTGT-TTGATGATGTTTGCATA 2558  
Db 11081 CAGATATTTTCTTGGGTGTATAGACCACAAAGACGTGTGTCTTGTATGATGTTTGCATA 11140  
Qy 2559 AAATTATACCTTAGTTTCTTGTATGTATGATGTAGATTTAGAGTATGAAATAG 2618  
Db 11141 AAATTATACCTTAGTTTCTTGTATGTATGATGTAGATTTAGAGTATGAAATAG 11200  
Qy 2619 TAGGAGGATTATTAACAAAGAACAGGCAAGAGTAGAATTAACCTCTTCTAATAC 2678  
Db 11201 TAGGAGGATTATTAACAAAGAACAGGCAAGAGTAGAATTAACCTCTTCTAATAC 11260  
Qy 2679 CTGTGCAAGTAGGCTTTTTCAGAACTCTACAA--CCCTACATAACTGGATAGTAGAA 2737  
Db 11261 CTGTGCAAGTAGGCTTTTTCAGAACTCTACAAACCCCAACAACTGGATAGTAGAA 11320  
Qy 2738 AAGCACCTCCCAAGGAGCGGTATCTTTGGAGTTTGAATCAGAGATAGAGCTAT 2797  
Db 11321 AAGCACCTCCCAAGGAGCGGTATCTTTGGAGTTTGAATCAGAGATAGAGCTAT 11380  
Qy 2798 AGCAATCTTCTATAGTAAACATAAAGAGCGCTGTTTATATATAGCAGTCAATTAAG 2857  
Db 11381 AGCAATCTTCTATAGTAAACATAAAGAGCGCTGTTTATATATAGCAGTCAATTAAG 11440  
Qy 2858 ATTTAAATAATACATCTTGCCTTCTTCTTACTCACAGATTTTCGAGAGGTAAATGTAAT 2917  
Db 11441 ATTTAAATAATACATCTTGCCTTCTTCTTACTCACAGATTTTCGAGAGGTAAATGTAAT 11500  
Qy 2918 GATC-CAGGAGGTGAGATCACTGCCTTTTATATGCGATTAAATTTGCATGAACAAAGTT 2976  
Db 11501 GATCAGAGGTGAGATCACTGCCTTTTATATGCGATTAAATTTGCATGAACAAAGTT 11559  
Qy 2977 TCCACAAATAACAGTAAATAAAGAAACATGTATTAGCACTTAATAAGCCAGGCGCTGG 3036  
Db 11560 TCCACAAATAACAGTAAATAAAGAAACATGTATTAGCACTTAATAAGCCAGGCTGT 11619  
Qy 3037 ACGAGGTGTATAGCTTTTCAATCCATGAACCTGGTAACTGGTACGATCTCTCTATT 3096  
Db 11620 ACGAGGTGTATAGCTTTTCAATCCATGAACCTGGTAACTGGTACGATCTCTCTATT 11679  
Qy 3097 GGACATGTGAGAAACCAATGGAGTTTCATAACAGTAGAGTTAAAAATTACTCTTCATA 3156  
Db 11680 GGACATGTGAGAAACCAATGGAGTTTCATAACAGTAGAGTTAAAAATTACTCTTCATA 11739  
Qy 3157 TATTATATGCTCAATCTCACAGACATCTCTGCTACCAAAAGCTATCATATCTAGA 3213  
Db 11740 TATTATATGCTCAATCTCACAGACATCTCTGCTACCAAAAGCTATCATATCTAGA 11796

ID AAC61382 standard; cDNA; 1579 BP.  
XX AC AAC61382;  
XX 19-FEB-2001 (first entry)  
XX cDNA sequence of androgen-related, prostate-specific gene PCGEM1.  
DE Androgen-regulated gene; prostate specific gene; PCGEM1; prostate cancer;  
KW prostate cancer gene expression marker 1; prostate related disease;  
KW benign prostatic hyperplasia; ss.  
XX OS Homo sapiens.  
XX PN WO2000058470-A1.  
XX PD 05-OCT-2000.  
XX 24-MAR-2000; 2000WO-US07906.  
XX PR 26-MAR-1999; 99US-0126469.  
XX (SRIK/) SRIKANTAN V.  
PA (ZOUZ/) ZOU Z.  
PA (MOUL/) MOUL J W.  
PA (SRIV/) SRIVASTAVA S.  
XX Srikantana V, Zou Z, Moul JW, Srivastava S;  
PI WPI; 2000-664926/64.  
DR Novel androgen-regulated prostate specific gene, prostate cancer gene  
XX expression marker, useful for detecting, diagnosing, preventing, and  
XX treating prostate cancer and other prostate related diseases -  
XX Claim 1; Fig 9; 75pp; English.  
CC The present sequence represents a cDNA fragment of an androgen-regulated,  
CC prostate specific gene PCGEM1 (prostate cancer gene expression marker 1).  
CC The PCGEM1 gene is over-expressed in prostate cancer. The PCGEM1  
CC polynucleotide is useful for detecting prostate cancer in a patient.  
CC The PCGEM1 promoter may be linked to cytotoxic gene, and be used  
CC for selectively killing prostate cancer cells. The PCGEM1 polynucleotide  
CC is also useful as marker of prostate cancer and other prostate related  
CC diseases, as targets for therapeutic intervention in prostate cancer  
CC and other prostate related diseases, in detection, diagnosis, prognosis,  
CC prevention, and treatment of prostate cancer (e.g. prostatic  
CC intraepithelial neoplasia (PIN), adenocarcinomas, nodular hyperplasia,  
CC and large duct carcinomas) and prostate related diseases (e.g. benign  
CC prostatic hyperplasia).  
XX SQ Sequence 1579 BP; 510 A; 275 C; 291 G; 503 T; 0 other;

Query Match 29.4%; Score 1531.6; DB 21; Length 1579;  
Best Local Similarity 99.7%; Pred. No. 2.8e-260;  
Matches 1545; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
Qy 216 GCAACTTCTCTAATTGGGAAATAGTTAAGCAGATTCTAGAGCTGAATAAATGT 275  
Db 15 GCAACTTCTCTAATTGGGAAATAGTTAAGCAGATTCTAGAGCTGAATAAATGT 74  
Qy 276 ACTACGAGATGCACCTGAGCTCAAGTGCACCTTATCAAGTGAATGGAGTCTTGCCTGT 335  
Db 75 ACTTCGAGATGCATGGGACTCAAGTGCACCTTATCAAGTGAATGGAGTCTTGCCTGT 134  
Qy 336 CTCGAAGCTGGAGCCCAATGTGTGATCTTGGCTCACTTGCACCTCCACCTCCAGGTT 395  
Db 135 CTCGAAGCTGGAGCCCAATGTGTGATCTTGGCTCACTTGCACCTCCACCTCCAGGTT 194  
Qy 396 CAAAGTTTCTCTGCTCAGCCCTCAGCCCTCCCAAGTAACTGGGATTACAGAGCTTTGGGCAATT 455  
Db 195 CAAAGTTTCTCTGCTCAGCCCTCAGCCCTCCCAAGTAACTGGGATTACAGAGCTTTGGTGCATT 254

```
Qy 456 TGACACTCATGATATCAGCCAAAGTGGAACTAAAAACAGCTCCTCGGAAGAGGACTATGA 515
Db 255 TGACACTTCATGATATCAGCCAAAGTGGAACTAAAAACAGCTCCTCGGAAGAGGACTATGA 314
Qy 516 CATCATCAGGTGGAGTCTCCAGGGACAGGGACCCCTTTGGGAAAAGGACTAGAAAAGTGT 575
Db 315 CATCATCAGGTGGAGTCTCCAGGGACAGGGACCCCTTTGGGAAAAGGACTAGAAAAGTGT 374
Qy 576 GAAATCATATAGTCTTCGATAGAAATCTCTGCTCTGTGTAAGAGCATTTTCATATTTACA 635
Db 375 GAAATCATATAGTCTTCGATAGAAATCTCTGCTCTGTGTAAGAGCATTTTCATATTTACA 434
Qy 636 AGACACAGGCTACTCTCCTAGGGCAGCAAAAAGTGGCAACAGCGCAAGCAGAGGAAAAGAG 695
Db 435 AGACACAGGCTACTCTCCTAGGGCAGCAAAAAGTGGCAACAGCGCAAGCAGAGGAAAAGAG 494
Qy 696 ATCATGAGGATTTTCAGATGCGACTGCTTTTTCATATATTTCTCAATGCCGTATGTTGG 755
Db 495 ATCATGAGGATTTTCAGATGCGACTGCTTTTTCATATATTTCTCAATGCCGTATGTTGG 554
Qy 756 TTTTATTTTGGCCCAAGCATACAAATCTGCTCAAGAAAAAATCTGGAAAAACAAGG 815
Db 555 TTTTATTTTGGCCCAAGCATACAAATCTGCTCAAGAAAAAATCTGGAAAAACAAGG 614
Qy 816 TGCCTTTGCCAATGTTATGTTTCTTTTGACAGCCCTGAGATTTCTGAGGGGAATTCAC 875
Db 615 TGCCTTTGCCAATGTTATGTTTCTTTTGACAGCCCTGAGATTTCTGAGGGGAATTCAC 674
Qy 876 ATAAATGGGATCAGGTCATTCATTTAGTTGTGTGCAAAATATGATTTTAAAGATACACCT 935
Db 675 ATAAATGGGATCAGGTCATTCATTTAGTTGTGTGCAAAATATGATTTTAAAGATACACCT 734
Qy 936 TTGCAGAGAGATGCTTTCTTAAGGTAGGCACGTGGAGCTAAGGTAAGGATTCCTT 995
Db 735 TTGCAGAGAGATGCTTTCTTAAGGTAGGCACGTGGAGCTAAGGTAAGGATTCCTT 794
Qy 996 CAAGATCAGTTAATCAAGAAAGTGCTTTTGGCATCTGAAATGCCCTTGTGCAAAAT 1055
Db 795 CAAGATCAGTTAATCAAGAAAGTGCTTTTGGCATCTGAAATGCCCTTGTGCAAAAT 854
Qy 1056 TGCTTATATGATTAAATTTACACTTAAATGGAACAACCTTTTAACTTACAGATGAACAAA 1115
Db 855 TGGTTATATGATTAAATTTACACTTAAATGGAACAACCTTTTAACTTACAGATGAACAAA 914
Qy 1116 -CCACAAAAGCAAAAATCAAAAGCCCTACCTATGATTTTCATATTTTCTGTGTAACCTG 1174
Db 915 CCCACAAAAGCAAAAATCAAAAGCCCTACCTATGATTTTCATATTTTCTGTGTAACCTG 974
Qy 1175 ATTAAGGATTCCTGCTGCTTTTGGGCATAAATGATAATGGAATATTTCCAGGTATGT 1234
Db 975 ATTAAGGATTCCTGCTGCTTTTGGGCATAAATGATAATGGAATATTTCCAGGTATGT 1034
Qy 1235 TTAATAAGAGGCCCTACTACAAATCTTACGATATCTTGGATATTTTAAATTCAGC 1294
Db 1035 TTAATAAGAGGCCCTACTACAAATCTTACGATATCTTGGATATTTTAAATTCAGC 1094
Qy 1295 TGGACATGTGCTAATCTTTTATATACATCTTTGCTAGATTTTCAATTTTAAAGTATG 1354
Db 1095 TGGACATGTGCTAATCTTTTATATACATCTTTGCTAGATTTTCAATTTTAAAGTATG 1154
Qy 1355 TGAATTTAGTTAATAGCTGTGCTGATCAATCAAAACATTAATCTTCTTAAATTTTGA 1414
Db 1155 TGAATTTAGTTAATAGCTGTGCTGATCAATCAAAACATTAATCTTCTTAAATTTTGA 1214
Qy 1415 CTATGAGGTCATAAATTCACAAAATATATCTACACATCAATATAGATGTTTTTCAT 1474
Db 1215 CTATGAGGTCATAAATTCACAAAATATATCTACACATCAATATAGATGTTTTTCAT 1274
Qy 1475 TATAATGCTCTCATCTTAAACAGAAATGCTCTTTTGTGATTGTTTTAGAAAACCTGAGAGTTT 1534
Db 1275 TATAATGCTCTCATCTTAAACAGAAATGCTCTTTTGTGATTGTTTTAGAAAACCTGAGAGTTT 1334
Qy 1535 TAATTCATAATTAATCTTGATCAAAAAAATGTGGGAACAATCCAGCATTAATTTGATGTGAT 1594
```

```
Db 1335 TAATTCATAATTAATCTTGATCAAAAAAATGTGGGAACAATCCAGCATTAATTTGATGTGAT 1394
Qy 1595 TGTTCATATGATGACATAGGAGTCTTAAGCTTTGGTGGCTTTGAAGTCTTTTGTACTTAGTCC 1654
Db 1395 TGTTCATATGATGACATAGGAGTCTTAAGCTTTGGTGGCTTTGAAGTCTTTTGTACTTAGTCC 1454
Qy 1655 CATCTTTAAAATTTACTACTTTTATATCTAAAGCATTTATGTTTTCATTTCAATTCATTTACATG 1714
Db 1455 CATCTTTAAAATTTACTACTTTTATATCTAAAGCATTTATGTTTTCATTTCAATTCATTTACATG 1514
Qy 1715 ATGCTAATTAATGCAATTAATAACAAATATTAAGATTTTCAAGATTCGAAATAGATAA 1764
Db 1515 ATGCTAATTAATGCAATTAATAACAAATATTAAGATTTTCAAGATTCGAAATAGATAA 1564

RESULT 3
AAC61381
ID AAC61381 standard; cdna; 1603 BP.
XX
AC AAC61381;
XX
DT 19-FEB-2001 (first entry)
XX
DE cdna sequence of androgen-related, prostate-specific gene PCGEM1.
XX
KW Androgen-regulated gene; prostate specific gene; PCGEM1; prostate cancer;
KW prostate cancer gene expression marker 1; prostate related disease;
KW benign prostatic hyperplasia; ss.
XX
OS Homo sapiens.
XX
PN WO200058470-A1.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US07906.
XX
PR 26-MAR-1999; 99US-0126469.
XX
PA (SRIK/) SRIKANTAN V.
PA (ZOUZ/) ZOU Z.
PA (MOUL/) MOUL J W.
PA (SRIV/) SRIVASTAVA S.
PI
XX Srikantan V, Zou Z, Moul JW, Srivastava S;
XX WPI; 2000-664926/64.
XX
PT Novel androgen-regulated prostate specific gene, prostate cancer gene
PT expression marker, useful for detecting, diagnosing, preventing, and
PT treating prostate cancer and other prostate related diseases
XX
PS Claim 1; Fig 8; 75pp; English.
XX
CC The present sequence represents a cDNA fragment of an androgen-regulated,
CC prostate specific gene PCGEM1 (prostate cancer gene expression marker 1).
CC The PCGEM1 gene is over-expressed in prostate cancer. The PCGEM1
CC polynucleotide is useful for detecting prostate cancer in a patient.
CC The PCGEM1 promoter may be linked to cytotoxic gene, and be used
CC for selectively killing prostate cancer cells. The PCGEM1 polynucleotide
CC is also useful as marker of prostate cancer and other prostate related
CC diseases, as targets for therapeutic intervention in prostate cancer
CC and other prostate related diseases, in detection, diagnosis, prognosis,
CC prevention, and treatment of prostate cancer (e.g. prostatic
CC intraepithelial neoplasia (PIN), adenocarcinomas, nodular hyperplasia,
CC and large duct carcinomas) and prostate related diseases (e.g. benign
CC prostatic hyperplasia).
SQ
Sequence 1603 BP; 531 A; 257 C; 293 G; 522 T; 0 other;
Query Match 27.9%; Score 1456.8; DB 21; Length 1603;
Best Local Similarity 92.8%; Pred. No. 4e-247;
```

Matches 1590; Conservative 0; Mismatches 2; Indels 122; Gaps 1;		
QY	51	AAGCAGCTCTGCGACCCAGTTTGGAACTGCGAGTTTAAAGTCAATAATGAATGAATA 110
Db	1	AAGGACCTCTGCGACCCAGTTTGGAACTGCGAGTTTAAAGTCAATAATGAATGAATA 60
QY	111	TGATAGCAAAAGGTGGAGGTTTTTAAAGAGCTATTATAGGTCCTCGACAGCATCTTTT 170
Db	61	TGATAGCAAAAGGTGGAGGTTTTTAAAGAGCTATTATAGGTCCTCGACAGCATCTTTT 120
QY	171	TCAATTAGGAGCAACCTTTTGGCCATGCGCGTAACCTGTGTCTGCAACTTCCTCTAAT 230
Db	121	TCAATTAGGAGCAACCTTTTGGCCATGCGCGTAACCTGTGTCTGCAACTTCCTCTAAT 180
QY	231	TGGAAATAGTTAAGCAGATTTCATAGAGCTCAATGATAAAATTTGTAACAGATGCACT 290
Db	181	TGGAAATAGTTAAGCAGATTTCATAGAGCTCAATGATAAAATTTGTAACAGATGCACT 240
QY	291	GGGACTCAACGTGACCTTATCAAGTGGAGTGGAGTCTTGGCCCTGTCTCCAAGGCTGGAGC 350
Db	241	GGGACTCAACGTGACCTTATCAAGTGGAGTGGAGTCTTGGCCCTGTCTCCAAGGCTGGAGC 266
QY	351	CCAATGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCAGGTTCAAACTGTTCTCCTG 410
Db	267	----- 266
QY	411	CCTCAGCCTCCCAAGTAACCTGGGATTACAGCAGGCTTGGTGCATTTGACATTCATGATA 470
Db	267	-----AGCAGGCTTGGTGCATTTGACATTCATGATA 298
QY	471	TCAGCAAAAGTGGAACTAAAAACAGCTCCTGGAAGAGGACTATGACATCATCAGTTGGG 530
Db	299	TCATCCAAAGTGGAACTAAAAACAGCTCCTGGAAGAGGACTATGACATCATCAGTTGGG 358
QY	531	AGTCTCCAGGACAGCGGACCCCTTTGGAAAAGGACTAGAAAAGTGAATCTATTAGTCT 590
Db	359	AGTCTCCAGGACAGCGGACCCCTTTGGAAAAGGACTAGAAAAGTGAATCTATTAGTCT 418
QY	591	TCGATATGAATCTCTGTCTCTGTAAAGCATTTTCAATGCGGTAATTTACAAGACAGCCCTACT 650
Db	419	TCGATATGAATCTCTGTCTCTGTAAAGCATTTTCAATGCGGTAATTTACAAGACAGCCCTACT 478
QY	651	CCTAGGCGAGCAAAAGTGGCAACAGGCAAGCAGAGGGAAGAAAGATCATGAGGCATTTTC 710
Db	479	CCTAGGCGAGCAAAAGTGGCAACAGGCAAGCAGAGGGAAGAAAGATCATGAGGCATTTTC 538
QY	711	AGAGTGCACTGCTCTTTTCATATATTTCTCAATGCGGTAATTTGGTTTATTTGGCCAA 770
Db	539	AGAGTGCACTGCTCTTTTCATATATTTCTCAATGCGGTAATTTGGTTTATTTGGCCAA 598
QY	771	GCATAACAATCTGCTCAAGAAAAAAATCTGGAGAAAAACAAAGTGCCCTTTGCCAATGT 830
Db	599	GCATAACAATCTGCTCAAGAAAAAAATCTGGAGAAAAACAAAGTGCCCTTTGCCAATGT 658
QY	831	TATGTTCTTTTTCAGAACGCTGAGATTTCTGAGGGGAATTCACATAAATGGATCAGG 890
Db	659	TATGTTCTTTTTCAGAACGCTGAGATTTCTGAGGGGAATTCACATAAATGGATCAGG 718
QY	891	TCATTCAATTTACGTTGTGCAATATGATTTAAAGATACAACTTTGCAGAGAGCATGC 950
Db	719	TCATTCAATTTACGTTGTGCAATATGATTTAAAGATACAACTTTGCAGAGAGCATGC 778
QY	951	TTTCCTAAGGTCAGGACGTCAGGACTTAAGGTTAAGCATTTCTCAAGATCAGTTAATC 1010
Db	779	TTTCCTAAGGTCAGGACGTCAGGACTTAAGGTTAAGCATTTCTCAAGATCAGTTAATC 838
QY	1011	AAGAAAGTGCTCTTTTGCATTTCTGAAATGCGCCTTTGTTGCAAAATATTGGTTATATTGATTA 1070
Db	839	AAGAAAGTGCTCTTTTGCATTTCTGAAATGCGCCTTTGTTGCAAAATATTGGTTATATTGATTA 898
QY	1071	AATTTACACTTAAATGGAAACAACTTTTAACCTTACAGATGAACAAACCCACAAAAGCAAAA 1130
Db	899	AATTTACACTTAAATGGAAACAACTTTTAACCTTACAGATGAACAAACCCACAAAAGCAAAA 958

QY	1131	AATCAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAACCTGGATTAAAGGATTCCTGCG 1190
Db	959	AATCAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAACCTGGATTAAAGGATTCCTGCG 1018
QY	1191	TTGCTTTTGGGCATAAATGATAAATGGAATATTTCCAGGTATTTGTTTAAATGAGGCGCCA 1250
Db	1019	TTGCTTTTGGGCATAAATGATAAATGGAATATTTCCAGGTATTTGTTTAAATGAGGCGCCA 1078
QY	1251	TCTCAAAATCTTAGCAATACTTTGGGATAATTTCAAAATTCAGCTGGACATTTGCTAATTT 1310
Db	1079	TCTCAAAATCTTAGCAATACTTTGGGATAATTTCAAAATTCAGCTGGACATTTGCTAATTT 1138
QY	1311	GTGTTTTATATACATCTTTTGGCTAGAAATTTCAAAATTTAAAGTATGTGAATTTAGTTAAATTA 1370
Db	1139	GTGTTTTATATACATCTTTTGGCTAGAAATTTCAAAATTTAAAGTATGTGAATTTAGTTAAATTA 1198
QY	1371	GCTGCTGTGATCAATTTCAAAACATTAATCTTTCCATAAATTTTAGACTATGAAGGTCATAAAA 1430
Db	1199	GCTGCTGTGATCAATTTCAAAACATTAATCTTTCCATAAATTTTAGACTATGAAGGTCATAAAA 1258
QY	1431	TTCAACAAATATATCTACACATACAAATATATAGATTTGTTTTTCATTTAATGCTTTCATCT 1490
Db	1259	TTCAACAAATATATCTACACATACAAATATATAGATTTGTTTTTCATTTAATGCTTTCATCT 1318
QY	1491	TAAACAGAAATCTCTTTTGTGATTTTGTAGAAAACCTGAGAGCTTTTAAATTCATAATTAATCT 1550
Db	1319	TAAACAGAAATCTCTTTTGTGATTTTGTAGAAAACCTGAGAGCTTTTAAATTCATAATTAATCT 1378
QY	1551	GATCAAAAAATTTGGGAAACAATCCAGCATTTAATTTGTATGTGATTTTATGTACATA 1610
Db	1379	GATCAAAAAATTTGGGAAACAATCCAGCATTTAATTTGTATGTGATTTTATGTACATA 1438
QY	1611	AGGAGTCTTAAGCTTGGTGCCCTTGAAAGTCTTTTGTACTTGTAGTCCCATGTTTAAATTAAT 1670
Db	1439	AGGAGTCTTAAGCTTGGTGCCCTTGAAAGTCTTTTGTACTTGTAGTCCCATGTTTAAATTAAT 1498
QY	1671	ACTTTATATCTAAAGCATTTATGTTTTTCAATTTCAATTTACATGATGCTAATTTATGGCAA 1730
Db	1499	ACTTTATATCTAAAGCATTTATGTTTTTCAATTTCAATTTACATGATGCTAATTTATGGCAA 1558
QY	1731	TTATAACAAATATTAAAGATTTTCGAAATAGATAA 1764
Db	1559	TTATAACAAATATTAAAGATTTTCGAAATAGATAA 1592
RESULT 4		
ABV23954/c		
ID	ABV23954	standard; cDNA; 1804 BP.
AC	ABV23954;	
XX		
DT	16-SEP-2002	(first entry)
XX		
DE	Human prostate expression marker cDNA 23945.	
DE		
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
KW	pharmacogenomic marker; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200160860-A2.	
PD	23-AUG-2001.	
XX		
PF	20-FEB-2001; 2001WO-US051171.	
XX		
PR	17-FEB-2000; 2000US-183319P.	
PR	16-MAR-2000; 2000US-189862P.	
PR	25-MAY-2000; 2000US-207454P.	
PR	09-JUN-2000; 2000US-211314P.	
PR	18-JUL-2000; 2000US-219007P.	
PR	13-DEC-2000; 2000US-255281P.	

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PA Schlegel R, Endege WO, Monahan JE;  
XX PI WPI; 2001-662795/76.  
XX DR Novel isolated nucleic acid molecule associated with cancerous state of  
XX PT prostate cells and correlating with presence of prostate cancer, useful  
XX PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX PS Claim 1; Page 4429; 11750pp; English.  
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX SQ Sequence 1804 BP; 529 A; 384 C; 333 G; 544 T; 14 other;  
  
Query Match 25.1%; Score 1308; DB 23; Length 1804;  
Best Local Similarity 94.7%; Pred. No. 6e-221;  
Matches 1507; Conservative 0; Mismatches 65; Indels 20; Gaps 14;  
  
QY 175 TTAGGACGCAACCTTTTGGCCCTATGCGGTAACTGCTGCTGCAACTTCTCTAA-TGGG 1744  
DB 1801 TTAGGACGCAACCTTTTGGCCCTATGCGGTAACTGCTGCTGCAACTTCTCTAA-TGGG 1744  
  
QY 235 AAATAGTTAGCAGATTCATAGAGCTGAATGATTAATTTCTACTACGAGATGCTAGGGA 294  
DB 1743 AAATAGTTAGCAGATTCATAGAGCTGAATGATTAATTTCTACTACGAGATGCTAGGGA 1684  
  
QY 295 CTCACGTGACCTTATCAAGTGAGATGGAGCTGTGCGCTGTCTCCAGGCTGGAGGCCAA 354  
DB 1683 CTCACGTGACCTTATCAAGTGAGATGGAGCTGTGCGCTGTCTCCAGGCTGGAGGCCAA 1624  
  
QY 355 TGGTGTGATCTTGGCTACATGCAACCTCCACCTCCAGGTTCAACGTTTCTCCTGCGCTC 414  
DB 1623 TGGTGTGATCTTGGCTACATGCAACCTCCACCTCCAGGTTCAACGTTTCTCCTGCGTG 1564  
  
QY 415 AGCCTCCCAAGTAACCTGGGATACAGCAGGCTTGGTGCAATTCACACTTCATGATACAG 474  
DB 1563 GACTCAACGTGACCTTATCAAGTGAGCAGGCTTGGTGCAATTCACACTTCATGATACAG 1504  
  
QY 475 CCAAAGTGGAACTAAACAGCTCTCGGAAGAGACTATGACATCATCAGGTTGGGAGTC 534  
DB 1503 CCAAAGTGGAACTAAACAGCTCTCGGAAGAGACTATGACATCATCAGGTTGGGAGTC 1444  
  
QY 535 TCAGGGACAGCGGACCCCTTTGGAAAGGAGTGAAGTGTGAATCTATTAGTCTTCGA 594  
DB 1443 TCAGGGACAGCGGACCCCTTTGGAAAGGAGTGAAGTGTGAATCTATTAGTCTTCGA 1384  
  
QY 595 TATGAATCTCTGCTCTGTAAAGCATTTTCATATTTACAGACAGCGCTACTCCTA 654  
DB 1383 TATGAATCTCTGCTCTGTAAAGCATTTTCATATTTACAGACAGCGCTACTCCTA 1324  
  
QY 655 GGCAGCAAAAGTGGCAACAGCAAGCAGAGGGAAGAGATCATGAGGCATTTTCAGAG 714  
DB 1323 GGCAGCAAAAGTGGCAACAGCAAGCAGAGGGAAGAGATCATGAGGCATTTTCAGAG 1264  
  
QY 715 TGCATGCTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTTATTTTGGCCAAAGCAT 774  
DB 715 TGCATGCTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTTATTTTGGCCAAAGCAT 774

DB 1263 TGCACCTGCTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTTATTTTGGCCAAAGCAT 1204  
QY 775 AACAACTGCTCAAGAAAAAAATCTGGAGAAACAAAGTGCCTTTTGGCAATGTTATG 834  
DB 1203 AACAACTGCTCA--AAAAAAATCTGGAGAAACAAAGTGCCTTTTGGCAATGTTATG 1146  
QY 835 TTTCTTTTGGACAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCAT 894  
DB 1145 TTTCTTTTGGACAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCAT 1086  
QY 895 TCATTTACGTTGTGTGCAAAATATGATTTAAAGATACAACTTTCAGAGAGAGATGCTTTC 954  
DB 1085 TCATTTACGTTGTGTGCAAAATATGATTTAAAGATACAACTTTCAGAGAGAGATGCTTTC 1026  
QY 955 CTAAAGGTAGCAGCTGGAGGACTAAGGTTAAAGCATTTCTCAAGATCAGTTAATCAAGA 1014  
DB 1025 CTAAAGGTAGCAGCTGGAGGACTAAGGTTAAAGCATTTCTCAAGATCAGTTAATCAAGA 966  
QY 1015 AAGGTGCTCTTTGCATTTCTGAAATGCCCTTTGTGCAAAATATTTGGTTATTTGATTAAT 1074  
DB 965 AAGGTGCTCTTTGCATTTCTGAAATGCCCTTTGTGCAAAATATTTGGTTATTTGATTAAT 906  
QY 1075 TACACTTAATGGAAACAACTTTAACTTACAGATGAACAAA-CCACAAAAAGCAAAAAT 1133  
DB 905 TACACTTAATGGAAACAACTTTAACTTACAGATGAACAAAACCCCAACAAAAAGCAAAAAT 846  
QY 1134 CAAAAGCCCTACCTATGATTTCTGTAACCTGATTTAAAGGATTCCTGCTTG 1193  
DB 845 CAAAAGCCCTACCTATGATTTCTGTAACCTGATTTAAAGGATTCCTGCTTG 786  
QY 1194 CTTTGGGCATAAATGGAATTTCCAGGATTTGTTTAAATAGGGGCCATCT 1253  
DB 785 CTTTGGGCATAAATGGAATTTCCAGGATTTGTTTAAATAGGGGCCATCT 726  
QY 1254 ACAAAATCTTAGCAATTTGGATTAATTTCTAAATTTACGTGGACATTTGCTTAATTTGTT 1313  
DB 725 ACAAAATCTTAGCAATTTGGATTAATTTCTAAATTTACGTGGACATTTGCTTAATTTGTT 666  
QY 1314 TTTTATATACATCTTCTGCTAGAAATTTCAAAATTTTAACTATGTAATTTAGTTAA-TTAGC 1372  
DB 665 TTTTATATACATCTTCTGCTAGAAATTTCAAAATTTTAACTATGTAATTTAGTTAA-TTAGC 606  
QY 1373 TGTGCT-GATCAATTTCAAAACATTTACTTTCTAAATTTTAGACTATGAAGGTCATAAAT 1431  
DB 605 TGTGCTGATCAATTTCAAAACATTTACTTTCTAAATTTTAGACTATGAAGGTCATAAAT 546  
QY 1432 TCAACAAATATATCTACACATACAAATTAATAGAT-TGTTTTTCATTAATGCTCTCATCT 1490  
DB 545 TCAACAAATATATCTACACATACAAATTAATAGATGTTGTTTTTCATTAATGCTCTCATCT 486  
QY 1491 TAACAGAAATGCTCTTCTGATTTGTTTTAGAAAACCTGAGAGTTTAAATTCATAATTTACTT 1550  
DB 485 TAACAGAAATGCTCTTCTGAT-TGTTTTAGAAAACCTGAGAGTTTAAATTCATAATTTACTT 427  
QY 1551 GATCAAAAAA--TTGTGGGAACAACTCCAGCATTAAT---TGTATGTTGATTTTATGTTATG 1605  
DB 426 GATCAAAAAATGTTGGGAACAACTCCAGCATTAATTTGATGCTGATTTGTTTATGTTATG 367  
QY 1606 ACATAAGGATCTTTAAGC--TTGGTGCCTTGAAGTCTTTTGTACTTTAGTCCCATGTTTAA 1663  
DB 366 ACATAAGGATCTTTAAGCCTGTTGGTGCCTTGAAGTCTTTTGTACTTTAGTCCCATGTTTAA 307  
QY 1664 AATTACTACTTTTATCTAAAGCATTTGTTTTCAATTTCAATTTACATGATGCTTAAT 1723  
DB 306 AATTACTA-TTTATATCTAAGCATTTATGTTTT--CAATCAATTAATGATGTTAAT 250  
QY 1724 ATGGCAATTAACAAATATTAAGATTTTCA 1755  
DB 249 AT-GCCATTATACAAATATTAAGGATCTCAA 219





Db 545 TCACAAATATATCTACACATACAAATATATAGATGTGTTTTTCATTTATATATGCTCTCATCT 486  
Qy 1491 TAACAGAAATGCTCTTTGTGATGTTTGTAGAAACTGAGAGTTTAAATTCATATTAATCTT 1550  
Db 485 TAACAGAAATGCTCTTTGTGA-TGGTTTTAGAAACTGAGAGTTTAAATTCATATTAATCTT 427  
Qy 1551 GATCAAAAAA--TTGTGGGAACATCCAGCATTAAT--TGTATGTGATGTTTTTTATG 1605  
Db 426 GATCAAAAAATGTTGGGAACAATCCAGCATTAATTTGTATGCTGATTTGTTTATGTT 367  
Qy 1606 ACATAAGAGTCTTAAGC--TTGGTGCTTGAAGTCTTTTGTACTTATAGTCCAGTTTAA 1663  
Db 366 ACATAAGAGTCTTAAGCTGTGGTGCTTGAAGTCTTTTGTACTTATAGTCCAGTTTAA 307  
Qy 1664 AATTACTACTTATATCTAAGCATTTATGTTTTTCAATTTCAATTTACATGATCTAAT 1723  
Db 306 AATTACTA-TTATATCTAAGCATTTATGTTTTT--CAATTTCAATTTACATGATGTTAAT 250  
Qy 1724 ATGGCAATTAACAATAATTAAGATTTTGA 1755  
Db 249 AT-GCCATTATAACAATAATTAAGGATCTCAA 219  
RESULT 6  
ABV29837/c  
ID ABV29837 standard; cdna; 1804 BP.  
XX AC ABV29837;  
XX 16-SEP-2002 (first entry)  
XX Human prostate expression marker cdna 29828.  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX W0200160860-A2.  
XX 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US05171.  
XX 17-FEB-2000; 2000US-183319p.  
XX 16-MAR-2000; 2000US-189862p.  
XX 25-MAY-2000; 2000US-207454p.  
XX 09-JUN-2000; 2000US-211314p.  
XX 18-JUL-2000; 2000US-219007p.  
XX 13-DEC-2000; 2000US-255281p.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX Claim 1; Page 6426; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC

CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 1804 BP; 529 A; 384 C; 333 G; 544 T; 14 other;  
Query Match 25.18; Score 1308; DB 23; Length 1804;  
Best Local Similarity 94.7%; Pred. No. 6e-221;  
Matches 1507; Conservative 0; Mismatches 65; Indels 20; Gaps 14;  
Qy 175 TTAGGCAGCAACCTTTTGGCCCTATGCCGTAACCTGTCTGCAACTTCCCTCTAATTTGGG 234  
Db 1801 TTAGGCAGCAACCTTTTGGCCCTATGCCGTAACCTGTCTGCAACTTCCCTCTAATTTGGG 1744  
Qy 235 AATATAGTTAAGCAGATTCATAGAGCTGAATGATATAAATTTGTAACGAGATGCTAGGGA 294  
Db 1743 AATATAGTTAAGCAGATTCATAGAGCTGAATGATATAAATTTGTAACGAGATGCTAGGGA 1684  
Qy 295 CTCAACGTGACCTTATCAAGTGAGATGGAGTCTTGGCCCTGTCTCAAGGCTGGAGCCCAA 354  
Db 1683 CTCAACGTGACCTTATCAAGTGAGATGGAGTCTTGGCCCTGTCTCAAGGCTGGAGCCCAA 1624  
Qy 355 TGGTGTGATCTTGGCTCACCTGCAACCTCCAGGTTCAAACTTTCTCTGCTCCTC 414  
Db 1623 TGGTGTGATCTTGGCTCACCTGCAACCTCCAGGTTCAAACTTTCTCTGCTCCTC 1564  
Qy 415 AGCCTCCCAAGTAATCGGATTTACAGCAGGCTTGGTGCATTTGACATTCATGATATCAG 474  
Db 1563 GACTCAACGTGACCTTATCAAGTGAGCAGGCTTGGTGCATTTGACATTCATGATATCAG 1504  
Qy 475 CCNAAGTGGAACTTAAACAGCTCTGGAAGAGACTATGACATCATCAGGTTGGGAGTC 534  
Db 1503 CCNAAGTGGAACTTAAACAGCTCTGGAAGAGACTATGACATCATCAGGTTGGGAGTC 1444  
Qy 535 TCCAGGACAGCGGACCCCTTTGGAAAGGACTGAAAGTGTGAAATCTATTAGTCTTCGA 594  
Db 1443 TCCAGGACAGCGGACCCCTTTGGAAAGGACTGAAAGTGTGAAATCTATTAGTCTTCGA 1384  
Qy 595 TATGAAATCTCTGCTCTGTATAAAGCATTTTCATATTTACAAAGACAGGCTCTCTCTA 654  
Db 1383 TATGAAATCTCTGCTCTGTATAAAGCATTTTCATATTTACAAAGACAGGCTCTCTCTA 1324  
Qy 655 GGCAGCAAAAAGTGGCAACAGCAAGCAGAGGAAAGAGATCATGAGCATTCAGAG 714  
Db 1323 GGCAGCAAAAAGTGGCAACAGCAAGCAGAGGAAAGAGATCATGAGGCAATTCAGAG 1264  
Qy 715 TGCACCTGCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTTTATTTTGGCCCAAGCAT 774  
Db 1263 TGCACCTGCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTTTATTTTGGCCCAAGCAT 1204  
Qy 775 AACAACTGCTCAAGAAAAAATCTGGGAAAAAAGAGTGCCTTTGCCAATGTTATG 834  
Db 1203 AACAACTGCTCA--AAAAAATCTGGGAAAAAAGAGTGCCTTTGCCAATGTTATG 1146  
Qy 835 TTTCTTTTTCACAGCCCTCAGATTTCTGAGGGAATTCACATAAATGGGATCAGTCAT 894  
Db 1145 TTTCTTTTTCACAGCCCTCAGATTTCTGAGGGAATTCACATAAATGGGATCAGTCAT 1086  
Qy 895 TCATTTACGTTGTGTGCAAAATATGATTTAAAGATACAACTTTTCAGAGAGCAGCTTTC 954  
Db 1085 TCATTTACGTTGTGTGCAAAATATGATTTAAAGATACAACTTTTCAGAGAGCAGCTTTC 1026  
Qy 955 CTAAAGGTAGGACGCTGGAGACTAAGGTAAGCATTTCTCAAGATCAGTTAATCAAGA 1014  
Db 1025 CTAAAGGTAGGACGCTGGAGACTAAGGTAAGCATTTCTCAAGATCAGTTAATCAAGA 966  
Qy 1015 AAGTGTCTCTTGCATTTCTGAAATGCCCTTGTTCGAAATATGTTGTTATTTGATTAAT 1074  
Db 965 AAGTGTCTCTTGCATTTCTGAAATGCCCTTGTTCGAAATATGTTGTTATTTGATTAAT 906  
Qy 1075 TACACTTAATGGAACAACCTTTAACTTACAGATGAACAAA-CCACAAAAAGCAAAAAAT 1133



Db 905 TACACTTAATGGAACAACCTTTAACTTACAGATGAACAAACCCACAAAGCAAAAT 846  
QY 1134 CAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGTAAAGGATTCCTGCTG 1193  
Db 845 CAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGTAAAGGATTCCTGCTG 786  
QY 1194 CTTTGGGCATAAATGATAATGGAATATTTCCAGGTATTTTAAATGAGGCCATCT 1253  
Db 785 CTTTGGGCATAAATGATAATGGAATATTTCCAGGTATTTTAAATGAGGCCATCT 726  
QY 1254 ACAATCTTAGCAATATTTGGATAATTTCAAAATTCAGTGGACATTTGCTAATGTT 1313  
Db 725 ACAATCTTAGCAATATTTGGATAATTTCAAAATTCAGTGGACATTTGCTAATGTT 666  
QY 1314 TTTTATATACATCTTGTAGCAATTTCAAAATTTAAGTATGTAATTTAGTTAA--TTAGC 1372  
Db 665 TTTTATATACATCTTGTAGCAATTTCAAAATTTAAGTATGTAATTTAGTTAA--TTAGC 606  
QY 1373 TGTGCT-GATCAATTCAAAACATTTACTTTCTCTAAATTTTAGACTATGAAGTCAATAAT 1431  
Db 605 TGTGTTGCATCAATTCAAAACATTTACTTTCTCTAAATTTTAGACTATGAAGTCAATAAT 546  
QY 1432 TCACAAATATATCTACACATACAAATTTATAGAT-TGTTTTTCATTAATGTCTTCATCT 1490  
Db 545 TCACAAATATATCTACACATACAAATTTATAGATGTGTTTTTCATTAATGTCTTCATCT 486  
QY 1491 TAACAGAATCTCTTTGTGATTTGTTTGTAGAACTGAGCTTTTAAATTCATAATTAAT 1550  
Db 485 TAACAGAATGCTTTGTGATTTGTTTGTAGAACTGAGCTTTTAAATTCATAATTAAT 427  
QY 1551 GATCAAAAA--TTGTGGGAACAATCCAGCATTAAT---TGATGTGATTTGTTTTATGT 1605  
Db 426 GATCAAAAAATGTTGGGAACAATCCAGCATTAATTTGTATGTTGTTTATGTT 367  
QY 1606 ACATAAGAGCTCTTAAGC--TTGTGCTTGAAGTCTTTTGTACTTATGCCAGTTTAA 1663  
Db 366 ACATAAGAGCTCTTAAGCTGTGGTGTCTTGAAGTCTTTTGTACTTATGCCAGTTTAA 307  
QY 1664 AATTACTACTTTATATCTAAGCAATTTATGTTTTCATTAATTTACATGATGCTAAT 1723  
Db 306 AATTACTA-TTTATATCTAAGCAATTTATGTTTTCATTAATTTACATGATGCTAAT 250  
QY 1724 ATGCCAATTATAACAATATTAAGATTTTCA 1755  
Db 249 AT-GCCATTATAACAATATTAAGCATCTCA 219

RESULT 7

ABV29916/c  
ID ABV29916 standard; cDNA; 1804 BP.  
XX AC ABV29916;  
XX DT 16-SEP-2002 (first entry)  
XX DE Human prostate expression marker cDNA 29907.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX PN WO2001160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US05171.  
XX PR 17-FEB-2000; 2000US-183319P.  
XX PR 16-MAR-2000; 2000US-189862P.  
XX PR 25-MAY-2000; 2000US-207454P.  
XX PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -  
XX Claim 1; Page 6454; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 1804 BP; 529 A; 384 C; 333 G; 544 T; 14 other;

Query Match 25.1%; Score 1308; DB 23; Length 1804;  
Best Local Similarity 94.7%; Pred. No. 6e-221;  
Matches 1507; Conservative 0; Mismatches 65; Indels 20; Gaps 14;  
QY 175 TTAGGCACACCTTTTGGCCCTATGCGTAACCTGTCTGCAACTTCTCTAATTTGGG 234  
Db 1801 TTAGGCACACCTTTTGGCCCTATGCGTAACCTGTCTGCAACTTCTCTAATTTGGG 1744  
QY 235 AAATAGTTAAGCAGATTCATAGAGTGAATGATAAATTTGACAGAGTGCATGCGGA 294  
Db 1743 AAATAGTTAAGCAGATTCATAGAGTGAATGATAAATTTGACAGAGTGCATGCGGA 1684  
QY 295 CTCACGTGACCTTTATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 354  
Db 1683 CTCACGTGACCTTTATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1624  
QY 355 TGGTGTGATCTTGGCTCACTCAACCTCCAGTTCACAGTTCACAACTTTCTCTGCTC 414  
Db 1623 TGGTGTGATCTTGGCTCACTCAACCTCCAGTTCACAACTTTCTCTGCTGCG 1564  
QY 415 AGCCTCCCAAGTAACTGGGATTACAGAGGCTTGGTGCATTTGACATTCATGATCAG 474  
Db 1563 GACTCAACGTGACCTTATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1504  
QY 475 CCAGTGGAACTAAACACAGCTCTCTGGAAGAGACTATGACATCATCAGTTGGAGTCT 534  
Db 1503 CCAGTGGAACTAAACACAGCTCTCTGGAAGAGACTATGACATCATCAGTTGGAGTCT 1444  
QY 535 TCCAGGACAGGACCTTTTGGAAAGAGACTAGAAAGTGAATCTATTAGTCTTCGA 594  
Db 1443 TCCAGGACAGGACCTTTTGGAAAGAGACTAGAAAGTGAATCTATTAGTCTTCGA 1384  
QY 595 TATGAATTTCTCTCTGTAAAGCATTTTATATTTTACAGACAGGCTTACTCTTA 654  
Db 1383 TATGAATTTCTCTCTGTAAAGCATTTTATATTTTACAGACAGGCTTACTCTTA 1324  
QY 655 GGGCAGCAAAAGTGGCAACAGGCAAGCAGAGGGAAGAGATCATGAGGATTTTCAGAG 714  
Db 1323 GGGCAGCAAAAGTGGCAACAGGCAAGCAGAGGGAAGAGATCATGAGGATTTTCAGAG 1264

```
QY 715 TGCACTGCTTTTCATATATTTCTCAATGCGGTATGTTTGGTTTATTTTGGCCAAACAT 774
Db 1263 TGCACTGCTTTTCATATATTTCTCAATGCGGTATGTTTGGTTTATTTTGGCCAAACAT 1204
QY 775 AACAAATGCTCAAGAAACAAAATCTGGAGAAACAAAGGTGCTTTTGGCAATGTTATG 834
Db 1203 AACAAATGCTCAAGAAACAAAATCTGGAGAAACAAAGGTGCTTTTGGCAATGTTATG 1146
QY 835 TTCTCTTTTGACAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCAT 894
Db 1145 TTTCTCTTTTGACAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCAT 1086
QY 895 TCATTTAGCTGTGTCGAATATGATTTAAAGATACAACTTGCAGAGAGCATGCTTTC 954
Db 1085 TCATTTAGCTGTGTCGAATATGATTTAAAGATACAACTTGCAGAGAGCATGCTTTC 1026
QY 955 CTAAGGCTAGGCGCTGGAGGACTAAGGGTAAGGACTTCTCAAGATCAGTTAATCAAGA 1014
Db 1025 CTAAGGCTAGGCGCTGGAGGACTAAGGGTAAGGACTTCTCAAGATCAGTTAATCAAGA 966
QY 1015 AAGTGTCTTTGCAATCTGAAATGCCCTTGTTCGAAATATTTGGTTATATGATTAAT 1074
Db 965 AAGTGTCTTTGCAATCTGAAATGCCCTTGTTCGAAATATTTGGTTATATGATTAAT 906
QY 1075 TACACTTAATGGAACAACTTTAACTTACAGATGAACAAA-CCACAAAAGCAAAAAT 1133
Db 905 TACACTTAATGGAACAACTTTAACTTACAGATGAACAAACCCACAAAAGCAAAAAT 846
QY 1134 CAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGGATTAAGGATTCCTGCTTG 1193
Db 845 CAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGGATTAAGGATTCCTGCTTG 786
QY 1194 CTTTTGGGCAATATGATATGAATATTTCCAGGTATTTTAAATGAGGGCCCATCT 1253
Db 785 CTTTTGGGCAATATGATATGAATATTTCCAGGTATTTTAAATGAGGGCCCATCT 726
QY 1254 ACAAATCTTAGCAATACTTTGGATAATTTCAAATTCAGTGGACATTTGCTAATGTT 1313
Db 725 ACAAATCTTAGCAATACTTTGGATAATTTCAAATTCAGTGGACATTTGCTAATGTT 666
QY 1314 TTTTATATAGATCTTTGCTAGAAATTCAAATTTTAAGTATGTGAATTTAGTTAA-TPAGC 1372
Db 665 TTTTATATAGATCTTTGCTAGAAATTTCAAATTTTAAGTATGTGAATTTAGTTAA 606
QY 1373 TGTGCT-GATCAATTCAAAACATCTACTTTCCCTAAATTTTAGACTATGAGGTCATAAT 1431
Db 605 TGTGCTGATCAATTCAAAACATCTACTTTCCCTAAATTTTAGACTATGAGGTCATAAT 546
QY 1432 TCAACAAATATATCTACACATACAAATATATAGAT-TGTTTTTCATTTAATGCTTTCATCT 1490
Db 545 TCAACAAATATATCTACACATACAAATATATAGATGCTGTTTTTCATTTAATGCTTTCATCT 486
QY 1491 TAACAGAAATGCTTTTGTGATTTTGTAGAACTGAGAGTTTAAATTCATTAATTAAT 1550
Db 485 TAACAGAAATGCTTTTGTGAT-TGGTTTTAGAAAACCTGAGAGTTTAAATTCATTAATTAAT 427
QY 1551 GATCAAAAAA--TGTTGGGAACAATCCAGCATTAAT--TGATGTGATTTGTTTTATGT 1605
Db 426 GATCAAAAAAATGTTGGGAACAATCCAGCATTAATTTGATGATGTTGTTTTATGTT 367
QY 1606 ACATAAGGAGCTCTTAAGC--TTGGTGCTTTGAAGTCTTTTGTACTTAGTCCCATGTTTAA 1663
Db 366 ACATAAGGAGCTCTTAAGCTGTTGGTGCTTTGAGTCTTTTGTACTTAGTCCCATGTTTAA 307
QY 1664 AATTACTACTTTATATCTAAGCATTTATGTTTTTCAATTCATTAATTTACATGATCTAAT 1723
Db 306 AATTACTA-TTTATATCTAAGCATTTATGTTTTT--CAATTCATTAATTTACATGATCTAAT 250
QY 1724 ATGGCAATTAACAATAATTAAGATTTTGA 1755
Db 249 AT-GCCATTAACAATAATTAAGGATCTCAA 219
```

## RESULT 8

```
AAC61384
ID AAC61384 standard; cDNA; 1025 BP.
XX AAC61384;
AC AC
XX AC
DT 19-FEB-2001 (first entry)
XX
DE Probe for androgen-related, prostate-specific gene PCGEM1 CDNA.
XX
KW Androgen-regulated gene; prostate specific gene; PCGEM1; prostate cancer;
KW prostate cancer gene expression marker 1; prostate related disease;
KW benign prostatic hyperplasia; probe; ss.
XX
OS Homo sapiens.
XX
WO200058470-A1.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US07906.
XX
PR 26-MAR-1999; 99US-0126469.
XX
PA (SRIK/) SRIKANTAN V.
PA (ZOUZ/) ZOU Z.
PA (MOUL/) MOUL J W.
PA (SRIV/) SRIVASTAVA S.
XX
PI Srikantan V, Zou Z, Moul JW, Srivastava S;
XX
XX WPI; 2000-664926/64.
XX
PT Novel androgen-regulated prostate specific gene, prostate cancer gene
PT expression marker, useful for detecting, diagnosing, preventing, and
PT treating prostate cancer and other prostate related diseases
XX
PS Claim 5; Fig 12; 75pp; English.
XX
CC The present sequence is a probe for cDNA of an androgen-regulated,
CC prostate specific gene PCGEM1 (prostate cancer gene expression marker 1).
CC The PCGEM1 gene is over-expressed in prostate cancer. The PCGEM1
CC polynucleotide is useful for detecting prostate cancer in a patient.
CC The PCGEM1 promoter may be linked to cytotoxic gene, and be used
CC for selectively killing prostate cancer cells. The PCGEM1 polynucleotide
CC is also useful as marker of prostate cancer and other prostate related
CC diseases, as targets for therapeutic intervention in prostate cancer
CC and other prostate related diseases, in detection, diagnosis, prognosis,
CC prevention, and treatment of prostate cancer (e.g. prostatic
CC intraepithelial neoplasia (PIN), adenocarcinomas, nodular hyperplasia,
CC and large duct carcinomas) and prostate related diseases (e.g. benign
CC prostatic hyperplasia).
XX
SQ Sequence 1025 BP; 327 A; 184 C; 208 G; 306 T; 0 other;
Query Match 17.1%; Score 891.4; DB 21; Length 1025;
Best Local Similarity 89.3%; Pred. No. 1e-147;
Matches 1024; Conservative 0; Mismatches 1; Indels 122; Gaps 1;
QY 166 TTTTTCATTTAGGCAACACCTTTTGGCCCTATGCGGTACCTGTCTGCAACTTCCT 225
Db 1 TTTTTCATTTAGGCAACACCTTTTGGCCCTATGCGGTACCTGTCTGCAACTTCCT 60
QY 226 CTAATTTGGGAAATAGTTAAAGCAGATTCATAGAGCTGAATATAAATTTGACTACGAGAT 285
Db 61 CTAATTTGGGAAATAGTTAAAGCAGATTCATAGAGCTGAATATAAATTTGACTACGAGAT 120
QY 286 GCACTGGGACTCAACGTGACCTTATCAAGTGAGATGAGAGTTCGCCCTGTCTCCAAAGGCT 345
Db 121 GCACTGGGACTCAACGTGACCTTATCAAGTG----- 151
QY 346 GGAGCCCAATGGTGTGATCTTGGCTCACTGCAACCTCCACTCCCGAGTTCAAACGGTTTC 405
```

Db 152 ----- 151  
QY 406 TCCTGCTCAGCTCCCAAGTAACCTGGGATACACAGCGCTTGGTGCATTTCACACTTCA 465  
Db 152 -----AGCAGGCTTGGTGCATTTCACACTTCA 178  
QY 466 TGATATCAGCCAAAGTGAACCTAAACAGCTCCTGGAAGAGGACTATGACATCATCAGG 525  
Db 179 TGATATCATCAAGTGAACCTAAACAGCTCCTGGAAGAGGACTATGACATCATCAGG 238  
QY 526 TTGGAGTCTCCAGGACAGCGGCCCTTTGGAAAAGGACTAGAAAGTGTGAATCTATT 585  
Db 239 TTGGAGTCTCCAGGACAGCGGCCCTTTGGAAAAGGACTAGAAAGTGTGAATCTATT 298  
QY 586 AGTCTTCGATATGAATCTCTGCTCTGTAAGCACTTTCATATTTACAACACACAGGC 645  
Db 299 AGTCTTCGATATGAATCTCTGCTCTGTAAGCACTTTCATATTTACAACACACAGGC 358  
QY 646 CTACTCTAGGSCAGCAAAAAGTGGCAACAGCAGAGGAGGAAAAGAGATCATGAGGC 705  
Db 359 CTACTCTAGGSCACAAAAGTGGCAACAGCAGAGGAGGAAAAGAGATCATGAGGC 418  
QY 706 ATTCAGAGTGCAGTCTCTTTCATATATTTCTCAATGCCGTATGTTGGTTTATTTTG 765  
Db 419 ATTCAGAGTGCAGTCTCTTTCATATATTTCTCAATGCCGTATGTTGGTTTATTTTG 478  
QY 766 GCCAAGCATAACAATCTGCTCAAGAAAAAATACTGAGAAAAACAAGGTCCTTGC 825  
Db 479 GCCAAGCATAACAATCTGCTCAAGAAAAAATACTGAGAAAAACAAGGTCCTTGC 538  
QY 826 AATGTTATGTTCTTTTGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGA 885  
Db 539 AATGTTATGTTCTTTTGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGA 598  
QY 886 TCAGTCTATCTATTCAGTGTGTGCAAAATATGATTTAAAGATACAACTTTGCAGAGAG 945  
Db 599 TCAGTCTATCTATTCAGTGTGTGCAAAATATGATTTAAAGATACAACTTTGCAGAGAG 658  
QY 946 CATGCTTCTTAAGGTAGGACGCGGAGGACTAAGGGTAAAGCATTTCTCAAGATCAGT 1005  
Db 659 CATGCTTCTTAAGGTAGGACGCGGAGGACTAAGGGTAAAGCATTTCTCAAGATCAGT 718  
QY 1006 TAATCAAGAAAGTCTCTTTGCAATCTGAAATGCCCTTTGCAAAATATTTGTTATATT 1065  
Db 719 TAATCAAGAAAGTCTCTTTGCAATCTGAAATGCCCTTTGCAAAATATTTGTTATATT 778  
QY 1066 GATTAATTTACACTTAATGGAACAACTTTAACTTACAGATGAACAAACCCACAAAG 1125  
Db 779 GATTAATTTACACTTAATGGAACAACTTTAACTTACAGATGAACAAACCCACAAAG 838  
QY 1126 CAAAAATCAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGTAAAGGATT 1185  
Db 839 CAAAAATCAAAAGCCCTACCTATGATTTTCATATTTCTGTGTAAGTAAAGGATT 898  
QY 1186 CTGCTTGTCTTTGGGCATAAATGATGAAATATTTCCAGGTATTTGTTAAATGAGG 1245  
Db 899 CTGCTTGTCTTTGGGCATAAATGATGAAATATTTCCAGGTATTTGTTAAATGAGG 958  
QY 1246 GCCCATCTACAAATTTCTAGCAATATCTTGGATAATCTTAAATTCAGCTGGACATGTC 1305  
Db 959 GCCCATCTACAAATTTCTAGCAATATCTTGGATAATCTTAAATTCAGCTGGACATGTC 1018  
QY 1306 TAAATGTT 1312  
Db 1019 TAAATGTT 1025

RESULT 9  
ABZ80229/c  
ID ABZ80229 standard; DNA; 249999 BP.  
XX  
AC ABZ80229;  
XX

DT 02-JUN-2003 (first entry)  
XX Human tramdorin gene region genomic DNA SEQ ID NO:26.  
DE Neuroprotective; nootropic; cerebroprotective; analgesic; gene therapy;  
XX central nervous system disorder; CNS disorder; multiple sclerosis;  
KW nerve injury; neuropathic pain; stroke; trauma; non-CNS disorder; tram;  
KW tramdorin; human; chromosome 5; gene; ds.  
XX Homo sapiens.  
OS WO2003016502-A2.  
XX 27-FEB-2003.  
XX 21-AUG-2002; 2002WO-US26637.  
PF 21-AUG-2001; 2001US-313907P.  
PR 21-AUG-2002; 2002US-0225810.  
XX (MCLA-) MCLAUGHLIN RES INST.  
XX Bermingham JR;  
XX WPI; 2003-278567/27.  
XX New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1,  
PT mouse tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd  
PT 3 or rat tramd 1, useful for treating CNS, e.g. stroke, multiple  
PT sclerosis, trauma, neuropathic pain -  
XX Example 6; Fig 9; 177pp; English.  
XX The present invention describes an isolated nucleic acid sequence  
CC comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse  
CC tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, or  
CC the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is  
CC located to chromosome 11, whereas human tramd 1 is located to chromosome  
CC 5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic  
CC and cerebroprotective activities, and can be used in gene therapy. The  
CC nucleic acid sequences are useful for diagnosing and treating central  
CC nervous system (CNS) disorders such as multiple sclerosis, nerve injury,  
CC neuropathic pain, stroke or trauma, and non-CNS disorders. The present  
CC sequence represents the genomic sequence of the human tramdorin gene  
CC region, which is given in the exemplification of the present invention.  
XX Sequence 249999 BP; 75050 A; 54012 C; 51931 G; 67805 T; 1201 other;  
SQ  
Query Match 8.9%; Score 462; DB 25; Length 249999;  
Best Local Similarity 84.7%; Pred No. 5, 7e-72;  
Matches 574; Conservative 0; Mismatches 80; Indels 24; Gaps 4;  
QY 3363 AGTGTGCCCGGAATTTATTTCTTGTGGTTCCTTGTGTCTGTGCTGCTCAAGAATGAA 3422  
Db 109978 AGTGTGTCCGGAATTTGTTCTTCTTGTGGTTCCTTGTGTCTTGTGCTCAAGAATGAA 109919  
QY 3423 ACCGACAGCCCTTGAGGTGAGTGTACAGTCTTAAAGATGTTGTTTCAGAGTTTGTTC 3482  
Db 109918 GCCACGGCCCTCGCGGTGAGTGTACAGTCTTAAAGATGTTGTTTCGCGAGTTTGTTC 109859  
QY 3483 CTTTCAGATGTTTCAGATGTTTCGCGAGTTCCTCCCTTATGTTGAGTTCGTTCCGCTGA 3542  
Db 109858 CTTTCAGATGTTTCAGATGTTTCGCGAGTTCCTCCCTTCTGTTGGTTCGTTCTGCTGA 109799  
QY 3543 CTTTCAGATGTTTCAGATGTTTCGCGAGTTCCTCCCTTATGTTGAGTTCGTTTCAGAGTTC 3602  
Db 109798 CTTTCAGAGTGAAGTGCAGACCTTTGCGAGTGA--GTGTTACAGCTCTTAAAGTGGGCGC 109741  
QY 3603 GTCCAGAGTTTGTTCCTCCCGGTAGTTCGTTCTGCTGATGTCAGGAATGAAGC 3662  
Db 109740 GTCCAGAGTTTGTTCCTCCCGGTAGTTCGTTCTGCTGATGTCAGGAATGAAGC 109681  
QY 3663 TGCAGACCTCGCGGTAAAGTGTACAGCTCATAAAGGTAGTGCATAAAGTGCAGAGTGAAGC 3722



```
RESULT 11
ABQ89916
ID ABQ89916 standard; CDNA; 843 BP.
XX
AC ABQ89916;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 1172.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US47349.
XX
PR 07-DEC-2000; 2000US-254648P.
XX
PR 13-MAR-2001; 2001US-275688P.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Krivenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
WPI; 2002-557824/59.
XX
New genes and gene products isolated from human prostate, useful for
treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
cancer), or as vaccines for treating or preventing these diseases -
XX
Claim 1; SEQ ID NO 1172; 186pp + Sequence Listing; English.
XX
The invention relates to an isolated polynucleotide comprising any of
1477 sequences or its fragment, degenerate variant, antisense or
complement. The polynucleotides and gene products are useful for treating
or diagnosing tumor or cancer (e.g. prostate cancer, breast cancer,
lung cancer or medullary carcinoma) in a subject (e.g. cattle, dogs,
cats, rabbits, horse or human). The polynucleotides and polypeptides are
also useful as vaccines for treating or preventing these diseases. The
polynucleotides are useful for gene therapy. The present sequence is that
of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
electronically as sequences of the invention. However only 1271
polynucleotide sequences are given, whereas 1477 polynucleotides and 91
proteins are claimed.
XX
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequence.
XX
Sequence 843 BP; 214 A; 190 C; 171 G; 251 T; 17 other;
XX
Query Match 8.4%; Score 438.4; DB 24; Length 843;
Best Local Similarity 99.8%; Pred. No. 4.6e-68;
Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
1 GAAACTTTAAATATCCCTCAGTGTCTCTGTTAATTCATGTAGTGCCCAAGGCACACTCT 60
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
123 GAAACTTTAAATATCCCTCAGTGTCTCTGTTAATTCATGTAGTGCCCAAGGCACACTCT 182
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
61 GGCACCCAGTTTGGAGTCAGTGTCTTAAAGTCATAAATGAATGAATAATGATACAAA 120
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
183 GGCACCCAGTTTGGAGTCAGTGTCTTAAAGTCATAAATGAATGAATAATGATACAAA 242
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
121 GGTGAGCTTTTAAAGAGCTATTATAGTGCCCTGGACAGCATCTTTTTCATTAGGC 180
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
243 GGTGAGGTTTTTAAAGAGCTATTATAGTGCCCTGGACAGCATCTTTTTCATTAGGC 302
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

```
181 AGCAACCTTTTGGCCCTATGCCGTAACTGTGTCTGCAACTTCCTCTAATTGGGAAATAG 240
|||||
303 AGCAACCTTTTGGCCCTATGTGTAACCTGTGTCTGCAACTTCCTCTAATTGGGAAATAG 362
|||||
241 TTAACGAGATTCATAGAGCTGAATGATATAAATTGTACTACGAGATGCACCTGGACCTCAAC 300
|||||
363 TTAACGAGATTCATAGAGCTGAATGATATAAATTGTACTACGAGATGCACCTGGACCTCAAC 422
|||||
301 GTGACCTTTATCAAGTGAGATGGAGTCTTTGCCCTGTCTCCAAAGGCTGGAGCCCAATGGTGT 360
|||||
423 GTGACCTTTATCAAGTGAGATGGAGTCTTTGCCCTGTCTCCAAAGGCTGGAGCCCAATGGTGT 482
|||||
361 GATCTTGGCTCACTGCAACCTCCACCTCCCAAGGTTCAAAACGTTTCTCTGCTGCATCAGCTC 420
|||||
483 GATCTTGGCTCACTGCAACCTCCACCTCCCAAGGTTCAAAACGTTTCTCTGCTGCATCAGCTC 542
|||||
421 CCAAGTAAGTGGGATTACAG 440
|||||
543 CCAAGTAAGTGGGATTACAG 562
|||||
RESULT 12
ABN95882/c
ID ABN95882 standard; DNA; 33206 BP.
XX
AC ABN95882;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2380 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
WPI; 2002-426119/45.
XX
Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
Claim 1; SEQ ID NO 2380; 298pp; English.
XX
The invention relates to a novel method for diagnosing and detecting the
progression of liver cancer, hepatocellular carcinoma or metastatic liver
tumour in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN97455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
```

XX	SQ	Sequence	33206 BP; 9901 A; 5689 C; 5955 G; 11661 T; 0 other;
		Query Match	8.1%; Score 421.2; DB 24; Length 33206;
		Best Local Similarity	80.3%; Pred. No. 7e-65;
		Matches	567; Conservative 0; Mismatches 108; Indels 31; Gaps 5;
Qy	3356	TCATATAAGTGTGCCCGGAATTAATTTCTTGTGGTGGTCTTGGTCTCGCTGACTCAAA	3415
Db	1173	TCATGCGAGTGTGCTCGGAATGGTTCCTTCCAATGGTTCCTGGTCTTGCCTGACTTCAA	1114
Qy	3416	GAATGAACCGCAGACCCCTTGAGTGAGTGTACAGTTCCTAAAGATGGTGTTCAGAG	3475
Db	1113	GAATGAAGCTGCGGACCTTTCGCGTGAGCATTTACAGTTTTAAAGATGGTGTGTCGGAG	1054
Qy	3476	TTTGTTCCTTTCAGATGTTACAGATGTGTCGCGAGTTTCTCCCTATGGTGTAGTTCGTGGTC	3535
Db	1053	TTTGTTCCTTTCAGATGTTACAGATGTGTCGAGATTTCTTCCTTGTGGGTTCCTCGTGC	995
Qy	3536	TCGCTGACTTCAACAATGAACCGCGACACCTTTGCAGTGAGTGTGACAGTTCCTAAAG	3595
Db	994	TTGCTGACTTCAGGAGTGAACCGCAGACCTTCACAGTGA - GTGTTCACAGCTCTAAAG	937
Qy	3596	GCAGTGGGTCCAGAGTTGTTGTTTCCTCCCGGTAGGTTTCGTGTGCTCGCTGATGTACAGA	3655
Db	936	ATGGTGGTCCGGAGTTGTTTCTTCCTCCGTGGTTCGAGGTCTTCGTGACTTCAGGA	877
Qy	3656	ATGAAGCTGCAGACCCCTCGCGGTAAAGTGTACAGCTCATAAAGGTAGTCAAAACCCAAAC	3715
Db	876	GTGAAGCCACAGACCTTCGCGGTGAGTGTACAGCTCATAAAGGTAGTGCAGACCCAAA	817
Qy	3716	AGTGAGCAGTAGCAAGATTTATTAAGAGCAAAAGAACAAAGCTTCCGCCACCATAGAA	3775
Db	816	AGAGCAGCAGCAAGATTTATTGCAAGAGACCAAGAACAAAGTTTCCACAGCTTGGAA	757
Qy	3776	ACGACACCAAAATTGGTTGCTGCTGCTG - - - TGGTAGCCAGCTTTTATTCCTTATTGG	3831
Db	756	GGGACACACACAGTTGCTGCGCTGGCTTCAGGTGGCCAGCTTTTATTCCTTATTGG	697
Qy	3832	CCACACCCATCCTGCTGATTGGCCCATTTTACAGAACTCATTTGGTGCATTTTATAG	3891
Db	696	TCCTCGCC - - - - - CACATCCCGCTCATTTGGGCCATTTTACAG	660
Qy	3892	CGTGTGATTGGTGCCTTTTACAGAGTGTGATTTGGTGCATTTACAACTCTTTAGCTAG	3951
Db	659	AGTGTGTTGGTCTGTTTTTACAAAGTGTGATTTGGTGCATTTTACAACTTTAGCTAG	600
Qy	3952	ACACAGAGTGTGATTGGTGCCTTTTAAATCCTTTTAGCTAGACACAAAAGTTCCTCAAGT	4011
Db	599	ACACAGAGCCTGATTGGTGTGTTTACAAATCCTTTAGCTAGACAGAAAAGTTCCTCCAAGT	540
Qy	4012	CCCCACCCAAACCA - GAAGCTCGCGTGGCTTACCTCTCGTAAAGG	4056
Db	539	CCCATATTCAGCCAGGAAGTCCAGTGGTTCACCTCTCAGCAGGA	494

RESULT 13  
AAC74309

ID AAC74309 standard; cDNA: 1761 bp.

AAC74309:

DT 02-FEB-2001 (first entry)

Human secreted protein gene 30 SEQ ID NO:40.

Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
carcinoprotective; nootropic; neuroprotective; antibacterial; virucide;
fungicide; ophthalmological; vulnary; gene therapy; angiogenesis;
autoimmune disease; hyperpliferative disorder; infection; skin aging;
wound healing; cardiovascular disorder; cerebrovascular disorder;
nervous system disorder; food additive; preservative; ss;

	Homo sapiens.	
XX	OS	
XX	PN	WO200057903-A2.
XX	PD	05-OCT-2000.
XX	PF	22-MAR-2000; 2000WO-USO7525.
XX	PR	26-MAR-1999; 99US-0126595.
XX	PT	22-DEC-1999; 99US-0171549.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	DR	WPI; 2000-594630/56.
XX	DR	P-PADB; AAB39339.
XX	PT	New nucleic acid molecules encoding 48 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -
XX	PS	Claim 1; Page 338; 395pp; English.
XX	CC	The polynucleotide sequences given in AAC74280 to AAC74327 encode the human secreted proteins given in AAB39310 to AAB39357. AAB393358 to AAB39400 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: antiarthritic; immunosuppressive; antirheumatic; anti-proliferative; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmologic; and vulnerary. The polynucleotides and polyptides can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders, cardiovascular disorders, cerebrovascular disorders, angiogenesis, nervous system disorders. Infections caused by bacteria, viruses and fungi and ocular disorders. The polyptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polyptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. AAC74271 to AAC74279 and AAB39309 represent sequences used in the exemplification of the present invention.
XX	SQ	Sequence 1761 BP; 409 A; 406 C; 439 G; 507 T; 0 other;
	Query Match	8.0%; Score 417.6; DB 2L; Length 1761;
	Best Local Similarity	82.7%; Pred. No. 2.3e-64;
	Matches	516; Conservative 0; Mismatches 99; Indels 9; Gaps 3;
Oy	3363	AGTGTCGCCGGAATTATTTTCTGTGGTGGTTCCTCGCTACTCCAAGAATAAA 3422
Dd	702	ATTATGTCTGGAATTGGTTCTCTCCGAGGGGTCTTTGGTCTCGCTCAAGAATAAA 761
Oy	3423	ACGCAGACCCTTWAGGTGAGTCTCAGATTCTTAAGATGGTGTTTCAGAGTTTTGTC 3482
Dd	762	GCCATGACCCCTCGCATGTAGTGTGCAGTCTTTAAAGATGGTGTGCCGGAGTTGTGT 821
Oy	3483	C TTCAGATGTTCAGATGTGTGC --- CGSAGTTTTCTCCCCTTATGTGTGAGTTCGTGGTCTCGC 3539
Dd	822	TTCAGATGTTCAGATGTGTGTCTCGAGAGTTTCTTCTCTGTGGGTCTGTGGTCTCGC 881
Oy	3540	TGACTTACAACAATGAAGCCGACAGACCTTTTGCATGTAGTGTGTGCACAGTTCTTTAAAGCAC 3599
Dd	882	TGACTTTCAGAGTGAAGCCACACACCTTCACTAGTA -- GCATTACAGCTCTTTAAAGGTGC 939

QY 3600 TCGCTCCAGAGTTGTTTCTCCCGTAGTTCGTGCTCGCTGATGTCAGGAATGA 3659  
DB 940 TCGCTCCAGAGTTGTTTCTCCCGTAGTTCGTGCTCGCTGATGTCAGGAATGA 999  
QY 3660 AGCTGCAGACCTTCGCGTAAGTGTACAGCTCATAAAGGTAGTGCACAAACCCAAACAGTG 3719  
DB 1000 AGCTGCAGACCTTCGCGTAAGTGTACAGCTCATAAAGGTAGTGCACAAACCCAAACAGTG 1059  
QY 3720 AGCAGTACAGAGTTATATGAAGACAAAGAAAGCTTCCCAACCATATGAACGG 3779  
DB 1060 AGCAGTACAGAGTTATATGAAGAGTATGAAGAAAGCTTCCCAACCATATGAACGG 1119  
QY 3780 ACCAGAAATGGTGTCTGC- ---TGCTGTGGTAGCCAGCTTTTATTCCTCTATTTCGGCCAC 3835  
DB 1120 ACCAGAAATGGTGTCTGC- ---TGCTGTGGTAGCCAGCTTTTATTCCTCTATTTCGGCCAC 1179  
QY 3836 ACCACATCTGCTGATGGGCCATTTTACAGAAATGCTGATGGTCCATTTATATGCGTG 3895  
DB 1180 GCCCTCATCTGCTGATGGTCCATTTTACAGAGGTGCTGATGGCCCATTTTACAGAGTG 1239  
QY 3896 CTGATTGCTGCGTTTTCACAGAGCTGATGGTGCATTTACATCTTTAGCTAGACAC 3955  
DB 1240 CTGATTGCTGCGTTTTCACAGAGCTGATGGTGCATTTTACATCTTTAGCTAGACAC 1299  
QY 3956 AGAGTCTGATGGTGCCTTTTATA 3979  
DB 1300 AGAGTCTGATGGTGCATTTTATA 1323

RESULT 14  
AAV83940/c  
ID AAV83940 standard; DNA; 80240 BP.  
XX  
AC AAV83940;  
XX  
XX 03-MAR-1999 (first entry)  
XX NC-contig derived from mardel(10) on chromosome 10q25.2.  
XX  
XX Yeast artificial chromosome; YAC; probe; eukaryotic chromosome;  
KW neocentromere; replication; extra-chromosomal element; segregation;  
KW cell division; artificial chromosome; gene therapy; mardel(10);  
KW human artificial chromosome; transgenic; chromosome 10; 10q25.2; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO9851790-A1.  
XX  
XX 19-NOV-1998.  
XX  
XX 13-MAY-1998; 98WO-AU00352.  
XX  
XX 26-AUG-1997; 97AU-0008791.  
PR 13-MAY-1997; 97AU-0006784.  
XX  
XX (AMRA-) AMRAD OPERATIONS PTY LTD.  
XX  
XX Cancilla MR, Choo K, Du Sart D;  
XX WPI; 1999-009773/01.  
XX  
XX New isolated nucleic acid comprising neocentromere sequences from  
PT eukaryotic chromosome - used to produce replicable, segregating  
PT artificial chromosomes that can carry large amounts of DNA for gene  
PT therapy  
XX  
XX Claim 9; Fig 16A; 540pp; English.  
XX  
XX The present sequence represents the NC-contig derived from a mutated  
CC human chromosome 10, 10q25.2 region. The sequence contains  
CC an unusual chromosomal marker referred to as mardel(10). The  
CC mardel(10) marker is mitotically stable and contains a functional

CC neocentromere at a location regarded as non-centromeric. This  
CC neocentromere maps to q25.2 on chromosome 10. The specification describes  
CC nucleic acid sequences derived from a eukaryotic chromosome, including a  
CC neocentromere or its functional derivative or hybrid, that are able, in  
CC a compatible cell, of replicating, acting as extra-chromosomal element  
CC and segregating during cell division. The sequences can be used to  
CC construct artificial chromosomes for use in gene therapy comprising a  
CC replicable, segregating nucleic acid that confers a specific phenotype  
CC on cells. Human artificial chromosomes can propagate in human cells and  
CC carry large amounts of DNA (e.g. therapeutic genes), and, being  
CC extra-chromosomal, they are not mutagenic. The artificial chromosomes  
CC are also useful for generation of transgenic plants and animals, in  
CC production of proteins and to make diagnostic reagents, e.g. for  
CC expression of cytokines, receptors and growth factors, or to increase  
CC the copy number of a gene in a cell. The constructs may also be  
CC used for functional and structural analysis of chromosomes.  
XX  
XX Sequence 80240 BP; 23102 A; 16537 C; 16747 G; 23846 T; 8 other;  
Query Match 7.8%; Score 404.6; DB 20; Length 80240;  
Best Local Similarity 79.7%; Pred No. 6.3e-62;  
Matches 551; Conservative 0; Mismatches 84; Indels 56; Gaps 4;  
QY 3363 AGTGTGCCCCGAATTTATTTCTTCGTGGGTCTTGTGCTCGCTGACTCCAAAGATGAA 3422  
DB 59807 ACTGTGCTCGGAATTTGTTCCTTCGGTGGTCTTGTGCTGACTTCAAGAATGAA 59748  
QY 3423 ACCGCAGACCTTGAGGTGAGTGCACAGTCTTAAAGATGGTGTTCAGAGTTTGTTC 3482  
DB 59747 GCCACAGACCTTGGCGTGGTGTACAGCTCTGAAAGATGGTGTCTGCTGAGTTTGTTC 59688  
QY 3483 CTTCAGATGTTTCAGATGTCGGGAGTTTCTCCCTTATGGTGTGCTGCTCGCTGA 3542  
DB 59687 CTTCAGATGTTTCAGATGTCGGGAGTTTCTCCCTTATGGTGTGCTGCTGCTGAC 59628  
QY 3543 CTTCAACATGAAGCCGACAGACCTTTGTCAGTGTGTCAGTGTCTTAAAGCGAGTGC 3602  
DB 59627 CTTCAAGAGTGAAGCTGCAGAGCTTCGCAGTGA--GTGTTACAGCTCTTAAAGCGAGCAC 59570  
QY 3603 GTCCAGAGTTGTTGTTCTCCCGTAGTGTGCTGCTGATGTCAGGAATGAAGC 3662  
DB 59569 ATCGGGAGTTCTTTGTTCTCCAGTGGGTGCTGCTGCT- GTTGAAGTTCAGGAATGAAGC 59511  
QY 3663 TGCAGACCTCGCGTAAAGTGTACAGCTCATAAAGTAGTGCACAAACCCAAACAGTGCAGC 3722  
DB 59510 CACAGACCTCGCAGTGTGTCAGCTCATAAAGTAGTGTGTCACACAAAGATGAGC 59451  
QY 3723 AGTAGCAAGATTTATTAAGAGACAAAGAGCTTCCCAACCATATGAAGCGAGC 3782  
DB 59450 AGCAGCAAGATTTATTTGAAGAGACAAAGAGCTTCCACAGTGTGGAAGGGGAGC 59391  
QY 3783 AGAATTGGTGTGCTGCTGCTG- ---TGGTAGCCAGCTTTTATTCCTTATTTGGCCACAGC 3838  
DB 59390 CGAGTGGGTGTCCACTGCTGGGTGGCCAGTTTTTACTCCCTTATTTGACCCCGCC 59331  
QY 3839 CACATCTGCTGATTTGGCCCATTTTACAGAAATGCTGATTTGGTCCATTTTATATGCTGCTG 3898  
DB 59330 CACATCTGCTGATTTGGTCCA----- 59310  
QY 3899 ATTGCTGCGTTTTTACAGAGTGTGATTTGGTGCATTTACAATCCCTTTAGCTAGACACAGA 3958  
DB 59309 -----TTTTACAGAGTGTGATTTGGTGCATTTACAATCCCTTTAGCTAGACACAGA 59260  
QY 3959 GTGCTGATTTGGTGCCTTTTATAATCCTTTAGCTAGACACAAAAGTCTACAAAGTCCCCACC 4018  
DB 59259 GTGCTGATTTGGTGCATTTTACAATCCTTTAGCTAGACACAAAAGTCTTCCAAAGTCCCCACC 59200  
QY 4019 CAACCCAGAAAGCTCGCTGGCTTCCACCTCTC 4049  
DB 59199 TGACCCAGAAAGCCAGCTGGCTTCCACCTCTC 59169





Result No.	Score	Query			ID	Description
		Match	Length	DB		
C 1	404.6	7.8	80246	3	US-09-078-294-4	Sequence 4, Appli
C 2	404.6	7.8	80595	3	US-09-078-294-3	Sequence 3, Appli
C 3	338.4	6.5	1043	4	US-09-432-576D-6	Sequence 6, Appli
C 4	311.4	6.0	1091	4	US-09-432-576D-5	Sequence 7, Appli
C 5	274.6	5.3	801	4	US-09-432-576D-7	Sequence 5, Appli
C 6	270.4	5.2	2834	3	US-09-305-384-6	Sequence 6, Appli
C 7	270.4	5.2	6235	3	US-09-305-384-5	Sequence 5, Appli
C 8	270.4	5.2	6235	4	US-09-535-160B-6	Sequence 6, Appli
C 9	270.4	5.2	6679	3	US-09-305-384-1	Sequence 1, Appli
C 10	270.4	5.2	6679	4	US-09-525-160B-5	Sequence 5, Appli
C 11	259	5.0	90541	4	US-09-759-359A-3	Sequence 3, Appli
C 12	255	4.9	1165	4	US-09-432-576D-25	Sequence 25, Appli
C 13	255	4.9	1831	4	US-08-432-576D-1	Sequence 1, Appli
C 14	205.6	3.9	2381	2	US-08-736-770-4	Sequence 4, Appli
C 15	176.2	3.4	19650	4	US-09-819-989-3	Sequence 3, Appli
C 16	174	3.3	46718	4	US-09-816-093-3	Sequence 3, Appli
C 17	172.8	3.3	9365	4	US-09-608-285A-4	Sequence 8, Appli
C 18	172.8	3.3	9365	4	US-09-350-836B-8	Sequence 8, Appli
C 19	172.8	3.3	9365	4	US-09-370-285-8	Sequence 8, Appli
C 20	172.8	3.3	9365	4	US-09-557-800C-8	Sequence 8, Appli
C 21	172.8	3.3	14747	4	US-09-608-285A-42	Sequence 42, Appli
C 22	172.8	3.3	14747	4	US-09-557-800C-42	Sequence 42, Appli
C 23	172.8	3.3	15977	4	US-09-608-285A-59	Sequence 59, Appli
C 24	172.6	3.3	35060	3	US-08-814-095-7	Sequence 7, Appli
C 25	170.8	3.3	43950	4	US-09-735-934A-3	Sequence 3, Appli
C 26	170.8	3.3	43950	4	US-10-060-332-3	Sequence 3, Appli
C 27	170.2	3.3	84495	4	US-09-797-906-3	Sequence 3, Appli















FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0145 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO.: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2381 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE: Consensus  
US-08-736-770-4

Query Match 3.9%; Score 205.6; DB 2; Length 2381;  
Best Local Similarity 80.28; Pred. No. 3.4e-35;  
Matches 279; Conservative 0; Mismatches 64; Indels 5; Gaps 3;  
Qy 3387 TGTGGTCTTCTGCTCGCTGACTCCAAAGATGAACCGCAGACCCCTTGAGGTGAGTGT 3446  
Db 344 TGTGGGTCTTGTCTCACTGACTTCAAGATGAAGCGCGGACCCCTAGCGTTGAGTGT 285  
Qy 3447 CACAGTCTTAAGATGGTGTGTCAGAGTTTGTTC-CTTCAGATGTTCAAGTGTGCCG 3505  
Db 284 CACGGTCTTAAAGTGGCGTGCAGGAGTTGTCTACTTCTGATGTTGCGATGTTG 225  
Qy 3506 GAGTTCTCCCTATGATGAGTGTGCTCGTGTGCTTCAAGATGAAGCGCGCAGACC 3565  
Db 224 GAGTTCTCCCTGTTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 166  
Qy 3566 TTTGCTAGTGTGTGACAGTCTTAAAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3625  
Db 165 TCGCGGTGAGTGTACAC--GCTCTTAAGCTGCGGCTCTGGAGTTGTTGTTCTCTCT 109  
Qy 3626 GGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3685  
Db 108 GGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 49  
Qy 3686 ACAGCTCATAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3733  
Db 48 ACAGCTCATAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1

RESULT 15  
US-09-819-989-3/c  
Sequence 3, Application US/09819989  
Patent No. 6482629  
GENERAL INFORMATION:  
APPLICANT: WEI, Ming-Hui et al.  
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001200  
CURRENT APPLICATION NUMBER: US/09/819,989  
CURRENT FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 19650  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(19650)  
OTHER INFORMATION: n = A,T,C or G  
US-09-819-989-3

Query Match 3.4%; Score 176.2; DB 4; Length 19650;  
Best Local Similarity 78.1%; Pred. No. 1.4e-28;  
Matches 257; Conservative 0; Mismatches 53; Indels 19; Gaps 3;  
Qy 1985 TTCAATTTTTTTTTTTTTTTTGGATAGAGATTTTGTCTTGTACACCCAGGTTGGAGTGCAT 2044  
Db 15437 TTTTTTTTTTTTTTTTTTTTGGATAGAGATTTTGTCTTGTACACCCAGGTTGGAGTGCAT 15379  
Qy 2045 GGCACGATCTCAGCTCACTGCAACCCCTGCTTCTCGTTCAAGTGAATTCCTCTGCTTC 2104  
Db 15378 GGCAGATCTCGGCTCACTGCAACCCCTGCTTCTCGTTCAAGTGAATTCCTCTGCTTC 15320  
Qy 2105 AGCTCCCAAGTAGCGGGGATTTACAGGTGCGCGCACCATCTGGTAAATTTTTTGTAT 2164  
Db 15319 AGCTCCCAAGTAGCTGGGATTTACAGACACACCCGCGGCTAAATTTTTTGTAT 15260  
Qy 2165 -----TTTCACCATGTTGGCCAGGCTGCTCGAACCCCTACCTCAGG 2207  
Db 15259 TTTTAGTAGAGACACAGGTTTTCACCATGTTGGCCAGGCTGCTCGAACCCCTACCTCAGG 15200  
Qy 2208 TGATCCCTCGCTCGGCTCTCAAAGTGTGGGATTTACAGGTGTAGCCACACGCCAG 2267  
Db 15199 GATTCACCCGCTTGGGCTCCCAAAGTGTGGGATTTACCCGCTGAGCCGCCCTGG 15140  
Qy 2268 CCCAATATCAGTTTTTTTCTTTTAAACACA 2296  
Db 15139 CTCCCGTTGTTTCTATATGCTGTCACA 15111

Search completed: September 23, 2003, 02:01:04  
Job time : 329 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 23:19:09 ; Search time 1780 Seconds  
(without alignments)  
7209.771 Million cell updates/sec

Title: US-09-996-956-1

Perfect score: 5217

Sequence: 1 gaaactttaaatccctc.....taaaaaaaaaaaaaaaa 5217

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1660708 seqs, 1229959015 residues

Total number of hits satisfying chosen parameters: 3321416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5217	100.0	5217	10	US-09-996-956-1
2	4957	95.0	5092	10	US-09-996-956-3
3	4758.4	91.2	32463	13	US-09-996-956-5
4	463	8.9	2057	13	US-10-027-632-100451
5	462	8.9	250000	12	US-10-225-810-26
6	421.2	8.1	33206	10	US-09-880-107-2380
7	404.6	7.8	80246	11	US-09-728-552-4
8	404.6	7.8	80595	11	US-09-728-552-3
9	394.8	7.6	650	13	US-10-027-632-226741
10	390.6	7.5	3212	11	US-09-764-891-8770
11	386.2	7.4	6138	11	US-09-764-891-8066
12	363.6	7.0	1691139	14	US-10-067-514-1
13	360.2	6.9	3037	13	US-10-027-632-116187
14	359.8	6.9	599	9	US-09-864-761-7986
15	359.4	6.9	142299	11	US-09-911-077A-14
16	350.8	6.7	861	13	US-10-027-632-154541

17	350.8	6.7	861	13	US-10-027-632-154542	Sequence 154542,
18	350.8	6.7	861	13	US-10-027-632-154543	Sequence 154543,
19	350.8	6.7	861	13	US-10-027-632-154544	Sequence 154544,
20	350	6.7	2044	13	US-10-027-632-99305	Sequence 99305, A
21	350	6.7	2044	13	US-10-027-632-99306	Sequence 99306, A
22	350	6.7	28818	10	US-09-764-877-2266	Sequence 2266, Ap
23	334.6	6.4	687	13	US-10-027-632-205124	Sequence 205124,
24	319	6.1	829	13	US-10-027-632-153282	Sequence 153282,
25	318.8	6.1	460	9	US-09-864-761-177	Sequence 177, App
26	314	6.0	1503841	9	US-09-795-668-1	Sequence 1, Appli
27	314	6.0	1503841	9	US-09-795-668-1	Sequence 1, Appli
28	314	6.0	1503841	10	US-09-946-807-1	Sequence 38, Appli
29	312	6.0	128779	15	US-10-081-327-38	Sequence 3, Appli
30	311	6.0	378361	11	US-09-901-136-3	Sequence 170030,
31	309.2	5.9	850	13	US-10-027-632-170030	Sequence 1878, Ap
32	308.2	5.9	6766	10	US-09-764-847-1878	Sequence 1878, Ap
33	308.2	5.9	6766	14	US-10-092-154-1878	Sequence 13, Appl
34	307.2	5.9	175590	11	US-09-911-077A-13	Sequence 3474, Ap
35	305.4	5.9	19172	10	US-09-764-877-3474	Sequence 6889, Ap
36	301.6	5.8	573	9	US-09-864-761-6889	Sequence 188551,
37	299.8	5.7	678	13	US-10-027-632-188551	Sequence 7928, Ap
38	298.4	5.7	500	9	US-09-864-761-7928	Sequence 13419, A
39	296.6	5.7	599	14	US-10-198-846-13419	Sequence 6749, Ap
40	296.6	5.7	864	14	US-10-198-846-6749	Sequence 3, Appli
41	294.8	5.7	10278	12	US-09-820-788-3	Sequence 200619,
42	294.6	5.6	632	13	US-10-027-632-200619	Sequence 182738,
43	293	5.6	603	13	US-10-027-632-182738	Sequence 182739,
44	293	5.6	603	13	US-10-027-632-182739	Sequence 1938, Ap
45	290.4	5.6	9180	12	US-10-311-455-1938	

#### ALIGNMENTS

#### RESULT 1

US-09-996-956-1  
; Sequence 1, Application US/09996956  
; Patent No. US20020155463A1  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies, Inc  
; TITLE OF INVENTION: Prostate Polynucleotides and Uses  
; FILE REFERENCE: 90 301 R1  
; CURRENT APPLICATION NUMBER: US/09/996.956  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 60/250,354  
; PRIOR FILING DATE: 2001-12-01  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 5217  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-996-956-1

Query Match	100.0%	Score 5217;	DB 10;	Length 5217;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 5217;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GAACCTTTAAATATCCCTCAGTCTCTCTGTTAATTCATGTAGTCCCAAGGCACCTCT	60	
Db	1	GAACCTTTAAATATCCCTCAGTCTCTCTGTTAATTCATGTAGTCCCAAGGCACCTCT	60	
Qy	61	GGCACCAGTTTGGACTGCAGTTTAAAGTCAATAATTGAATGAAATGATAGCAAA	120	
Db	61	GGCACCAGTTTGGACTGCAGTTTAAAGTCAATAATTGAATGAAATGATAGCAAA	120	
Qy	121	GGTGAGGTTTTTAAAGAGCTATTATAGTCCCTGGACAGCATCTTTTCAATTAGGC	180	
Db	121	GGTGAGGTTTTTAAAGAGCTATTATAGTCCCTGGACAGCATCTTTTCAATTAGGC	180	
Qy	181	AGCAACCTTTTGGCCCTATGCCGTAACCTGTCTGCAACTTCTCTAATTTGGGAATAG	240	
Db	181	AGCAACCTTTTGGCCCTATGCCGTAACCTGTCTGCAACTTCTCTAATTTGGGAATAG	240	

QY	241	TTAAGCAGATTCTAGAGCTGAATGATAAAATTGTACTACGAGATGCACCTGGGACTCAAC	300
Db	241		
QY	301	GTGACCTTATCAAGTGAGATGGAGCTTGGCCCTGTCTCCAAAGGCTGGAGCCCAATGGTGT	360
Db	301		
QY	361	GATCTTGGCTCACTGCAGCTCCACCTCCAGGTTCAACGGTTTCTCCTGCCCTCAGGCTC	420
Db	361		
QY	421	CCAAGTAACCTGGGATTACAGCAGGCTTGGTGCAATTCACACCTTCATGATATCAGCCAAAG	480
Db	421		
QY	481	TGGAACCTAAAACAGCTCCTGGAAGAGGACTATGACATCATCAGGTTGGAGTCTCCAGG	540
Db	481		
QY	541	GACAGCGGACCTTTGAAAGAGCTAGAAAGTGTAAATCTATTAGTCTTCGATATGAA	600
Db	541		
QY	601	ATTCTCTGCTCTGTAAGACATTTTCATATTTACAAGACACAGGCCCTACTCCTAGGGCAG	660
Db	601		
QY	661	CAAAAAGTGCAACAGCAAGCAAGGGAAGAGATCATGAGGCAATTCAGAGTGCACCT	720
Db	661		
QY	721	GTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTATTTTGGCCCAAGCATACAAT	780
Db	721		
QY	781	CTGCTCAAGAAAAAATCTGGAGAAAAAAGGTCGCTTTGCCAATGTTATGTTCTT	840
Db	781		
QY	841	TTTGACAGCCCTGAGATTCTGAGGGGAATTCACATAATGGGATCAGGTCAATCATTT	900
Db	841		
QY	901	ACGTTGTGCAAAATGATTTAAAGATACAACCTTTGCAGAGCATGCTTTCTTAAGG	960
Db	901		
QY	961	GTAGGCACGTGGAGGACTAAGGTAAGCAATCTTCAAGATCAGTTAATCAAGAAAGGTG	1020
Db	961		
QY	1021	CTCTTTGCATCTGAAATGCCCTTGTGCAAAATTTGGTTATATTGATTAATTTACACT	1080
Db	1021		
QY	1081	TAATGGAACAACCTTTAACTTACAGATGAACAAACCCACAAAGCAAAAAATCAAAAGC	1140
Db	1081		
QY	1141	CCTACCTATGATTTTCATATTTCTGTAACTGGATTAAGGATTCCTGCTCTTTTGG	1200
Db	1141		
QY	1201	GCATAAATGATAATGAAATTTCCAGGTTATTTGTAATTTGATTAATTTACAAAT	1260
Db	1201		
QY	1261	CTTAGCAATCTTTGGATTAATCTAAAAATTCAGCTGGACATTCGTCTAAATGTTTTTATA	1320
Db	1261		
QY	1321	TACATCTTTGCTAGAATTTCAAAATTTAAAGTATGTGAATTTAGTAAATTTAGCTGTGCTGA	1380
Db	1321		
QY	1381	TCAATTCAAAACAATTTACTTTTCCATAATTTTAGACTATGAAGGTTCATAAAATTCACAAAT	1440
Db	1381		
QY	1441	ATATCTACACATACAATATAGATGTTTTTCATTATAATGCTCTCATCTTAACAGAAAT	1500
Db	1441		
QY	1501	GTCTTTGTGATGTTTTTGAAGAACTGAGAGTTTTTAATTCATATTTACTTGATCAAAAAA	1560
Db	1501		
QY	1561	TTGTGGGAACAATCCAGCATTAATGTATGTGATGTTTTTATGTACATAGGAGTCTTA	1620
Db	1561		
QY	1621	AGCTTGTGCTCTGCAAGTCTTTTGTACTTGTAGTCCCATGTTTAAATTTACTACTTTATC	1680
Db	1621		
QY	1681	TAAAGCATTTATGTTTTTCAATTTACATGATCTAATTTATGSCAATTTATAACAAA	1740
Db	1681		
QY	1741	TATTAAGAATTCGAAATAGAATATGTGAATTTGTTACATACATAGAAATGAAAGTTCA	1800
Db	1741		
QY	1801	TTTTCGTAAGCAAGATCTCGGTCGAAGAGTGGTTTTGATTGAAAGATCACTAGATTAGT	1860
Db	1801		
QY	1861	AGAGGCAAGCTTTAGTCCCTAATCTACCTTAATAGCCATGTGTGTCACGTGAAGTC	1920
Db	1861		
QY	1921	AGTGAACCCATCTCATCTCTCTCATCTTTTTCATCTCTAAAATGAGGATATAATTTAA	1980
Db	1921		
QY	1981	GCCTTCATTTTTTTTTTTTTTTTGGATAGAGTTTTGCTCTGTACCCAGGTTGGAGTG	2040
Db	1981		
QY	2041	CAATGGCAGCATCTAGCTCACTGCAACCTCTGCTTCCGGTTCAAGTGAATTCCTCTG	2100
Db	2041		
QY	2101	CTTCAGCCCTCCCAAGTAGCCGGATTACAGTGCCGCCACCATCTGCTATTTTTTT	2160
Db	2101		
QY	2161	GTATTTTACCATGTTGGCCAGGCTGGTCTCGAACCCTACCTCAGGTGATCCCTCGCCT	2220
Db	2161		
QY	2221	CGCCCTCTCAAGTGTGGGATTACAGGTGTAGCCACACGCCCCAGCCCAATATCAGTT	2280
Db	2221		
QY	2281	TTTCTTTTTTAAACAAGGCTAACAAATCAAAATCTAGCTAGGGGAGAAAAAAT	2340
Db	2281		
QY	2341	AAGGCACTGTTTATGTAAACAGCTCTTGTGCAATTCAGTGGCAGACAAATAAACAGTA	2400
Db	2341		
QY	2401	AGAATCAATCCTTTTCATATATCCTTCTTGCAGAAATACATAAAATCCCAAAATGGCTAT	2460
Db	2401		



```
QY 4621 AAAGAGTTTGCAATATAAGATTCCCTTTCTTTGTTAAACATTTCTACCTTGGGGCTTCATTT 4680
DB 4621 AAAGAGTTTGCAATATAAGATTCCCTTTCTTTGTTAAACATTTCTACCTTGGGGCTTCATTT 4680
QY 4681 ATAATCAAAAGGAGTACTGTAACCTGTCAAAAAAAGCTACCTGTGACAAATATATTATGT 4740
DB 4681 ATAATCAAAAGGAGTACTGTAACCTGTCAAAAAAAGCTACCTGTGACAAATATATTATGT 4740
QY 4741 GATGGTTACCTGACGTAGGTGGCAATTAATAATAATTAATCAAGAAATCAAGCCG 4800
DB 4741 GATGGTTACCTGACGTAGGTGGCAATTAATAATAATTAATCAAGAAATCAAGCCG 4800
QY 4801 AGCAGAACTGTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAA 4860
DB 4801 AGCAGAACTGTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAA 4860
QY 4861 TAATCATATAAATPAGCCACTCAATCCAAAACATCACTGGGCGACTTGTACATATATAA 4920
DB 4861 TAATCATATAAATPAGCCACTCAATCCAAAACATCACTGGGCGACTTGTACATATATAA 4920
QY 4921 TCAGTGGAGATGTAATGAAGCACAAGCTTAAGTGAATGTCTAGAGAGCTAATGATTC 4980
DB 4921 TCAGTGGAGATGTAATGAAGCACAAGCTTAAGTGAATGTCTAGAGAGCTAATGATTC 4980
QY 4981 ATTTTATGGAATTTTACTTATTTTAAATGTCTATCCCTGACCATCTTGAACCTTTACTT 5040
DB 4981 ATTTTATGGAATTTTACTTATTTTAAATGTCTATCCCTGACCATCTTGAACCTTTACTT 5040
QY 5041 GAAGATTTATTTTTTTTAAATCACTGTTTATTAGATTTAGGTATCTGGTCTTTGTT 5100
DB 5041 GAAGATTTATTTTTTTTAAATCACTGTTTATTAGATTTAGGTATCTGGTCTTTGTT 5100
QY 5101 TTTCTTTTATCATGTATGATTTTATTTTATTTTATGTCAGTGTCCCTTAAGCTTCATCAA 5160
DB 5101 TTTCTTTTATCATGTATGATTTTATTTTATTTTATGTCAGTGTCCCTTAAGCTTCATCAA 5160
QY 5161 TGAGAAGAAATGTAATAATCCATTTATTTTACCCTTAATAAAAAA 5217
DB 5161 TGAGAAGAAATGTAATAATCCATTTATTTTACCCTTAATAAAAAA 5217
```

## RESULT 2

```
us-09-996-956-3
; Sequence 3, Application US/09996956
; Patent No. US20020155463A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Prostate Polynucleotides and Uses
; FILE REFERENCE: 9U 301 R1
; CURRENT APPLICATION NUMBER: US/09/996.956
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,354
; PRIOR FILING DATE: 2001-12-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5092
; TYPE: DNA
; ORGANISM: Homo sapiens
us-09-996-956-3
```

```
Query Match          95.0%; Score 4957; DB 10; Length 5092;
Best Local Similarity 97.68; Pred. No. 0;
Matches 5092; Conservative 0; Mismatches 0; Indels 125; Gaps 1;
```

```
QY 1 GAAACTTTAAATATATCCCTCAGTGCCTCTGTTAATTCATGGTAGTGCCTCCCAAGGCACTCT 60
DB 1 GAAACTTTAAATATATCCCTCAGTGCCTCTGTTAATTCATGGTAGTGCCTCCCAAGGCACTCT 60
QY 61 GGCACCCAGTTTTGGAACTGCAGTTTTTAAAGTCATAAATTGAATGAAATGATAGCAA 120
DB 61 GGCACCCAGTTTTGGAACTGCAGTTTTTAAAGTCATAAATTGAATGAAATGATAGCAA 120
```

```
QY 121 GGTGGAGGTTTTAAAGAGCTATTTATAGGTCCCTGGACAGCATCTTTTCAATTAGGC 180
DB 121 GGTGGAGGTTTTAAAGAGCTATTTATAGGTCCCTGGACAGCATCTTTTCAATTAGGC 180
QY 181 AGCAACCTTTTGGCCCTATGCCGTAACCTGTGCTGCAAACTTCCCTCTAAATGGGAAATAG 240
DB 181 AGCAACCTTTTGGCCCTATGCCGTAACCTGTGCTGCAAACTTCCCTCTAAATGGGAAATAG 240
QY 241 TTAAGCAGATTCATPAGAGCTGAATGATAAAATTTACTACGAGATGCACCTGGGACTCAAC 300
DB 241 TTAAGCAGATTCATPAGAGCTGAATGATAAAATTTACTACGAGATGCACCTGGGACTCAAC 300
QY 301 GTGACCTTTATCAAGCTGAGATGGAGCTCTTGCCCTGTCTCCRAGGCTGGAGGCCAATGGTGT 360
DB 301 GTGACCTTTATCAAGCTG----- 316
QY 361 GATCTTGGCTCACTGCAACCTCCACCTCCAGGTTCAAAACGTTTCTCTCTGCTCAGCCTC 420
DB 317 ----- 316
QY 421 CCAAGTAACCTGGGATTAACAGCAGGCTTGGTGCATTTGACACTTTCATGATATCAGCCAAAG 480
DB 317 -----AGGCTTGGTGCACTTGGACACTTTCATGATATCAGCCAAAG 355
QY 481 TGGAACTAAAACAGCTCCCTGGGAAGAGGACTATCACATCATCAGGTTGGGAGTCTCCAGG 540
DB 356 TGGAACTAAAACAGCTCCCTGGGAAGAGGACTATCACATCATCAGGTTGGGAGTCTCCAGG 415
QY 541 GACAGCGGACCCCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAA 600
DB 416 GACAGCGGACCCCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAA 475
QY 601 ATTCTCTGTCTCTGTAAAAGCATTTTCATATTTTACAACACACAGGCTTACTCCTTAGGCGAG 660
DB 476 ATTCTCTGTCTCTGTAAAAGCATTTTCATATTTTACAACACACAGGCTTACTCCTTAGGCGAG 535
QY 661 CAAAAGTGGCAACAGCAGCAAGGAAAGAGATCATGAGGCATTTTCAGAGTGCACCT 720
DB 536 CAAAAGTGGCAACAGCAGCAAGGAAAGAGATCATGAGGCATTTTCAGAGTGCACCT 595
QY 721 GTCTTTTCATATATTTCTCAATGCCGTATGTTGGTGTATTTTGGCCACAGCATCAACAAT 780
DB 596 GTCTTTTCATATATTTCTCAATGCCGTATGTTGGTGTATTTTGGCCACAGCATCAACAAT 655
QY 781 CTGCTCAAGAAAAAATCTGGAGAAACAAAGGTGCCCTTGGCCATGTTATGTTCTT 840
DB 656 CTGCTCAAGAAAAAATCTGGAGAAAAAACAAGGTGCCCTTGGCCATGTTATGTTCTT 715
QY 841 TTTGACAAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCATTCATTT 900
DB 716 TTTGACAAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCATTCATTT 775
QY 901 ACCTTTGTGCAAAATATGATTTTAAAGATACAACTTTTGCAGAGAGCATGCTTTCTTAAGG 960
DB 776 ACCTTTGTGCAAAATATGATTTTAAAGATACAACTTTTGCAGAGAGCATGCTTTCTTAAGG 835
QY 961 GTAGGCACGTGGAGGACTAAGGTAAGGCATTTCTCAAGATCAGTTAATCAAGAAAGGTG 1020
DB 836 GTAGGCACGTGGAGGACTAAGGTAAGGCATTTCTCAAGATCAGTTAATCAAGAAAGGTG 895
QY 1021 CTCTTTGCAATCTGAAATGCCCTTGTGCAAAATATTGGTTATATTGATTTAAATTTACACT 1080
DB 896 CTCTTTGCAATCTGAAATGCCCTTGTGCAAAATATTGGTTATATTGATTTAAATTTACACT 955
QY 1081 TAATGGAACCAACTTTTAACTTACAGATGAACAAACCCACAAAGCAAAAAATCAAAAGC 1140
DB 956 TAATGGAACCAACTTTTAACTTACAGATGAACAAACCCACAAAGCAAAAAATCAAAAGC 1015
QY 1141 CCTACCTATGATTTTCATATTTTCTGTGTAAGTAAAGGATTCCTGCTTGGTTTGG 1200
DB 1016 CCTACCTATGATTTTCATATTTTCTGTGTAAGTAAAGGATTCCTGCTTGGTTTGG 1075
```

Qy	1201	GCATAAATGATAATGGAAATATTTCCAGGTATTTGTTTAAATAGAGGGCCCACTACAAATTT	1261
Db	1076	GCATAAATGATAATGGAAATATTTCCAGGTATTTGTTTAAATAGAGGGCCCACTACAAATTT	1135
Qy	1261	CTTAGCAATACTTTGGATAATTCCTAAATTTCCAGCTGGACATGCTCTAAATGTTTATTATA	1320
Db	1136	CTTAGCAATACTTTGGATAATTCCTAAATTTCCAGCTGGACATGCTCTAAATGTTTATTATA	1195
Qy	1321	TACATCTTTGCTAGAAATTTCAAATTTTAAAGTATGTGAATTTAGTTAAATAGCTGTGCTGA	1380
Db	1196	TACATCTTTGCTAGAAATTTCAAATTTTAAAGTATGTGAATTTAGTTAAATAGCTGTGCTGA	1255
Qy	1381	TCAATTCAAAAACAATTACTTTCTTAAATTTTAGACTATGAAGTCAATAAATCAACAAT	1440
Db	1256	TCAATTCAAAAACAATTACTTTCTTAAATTTTAGACTATGAAGTCAATAAATCAACAAT	1315
Qy	1441	ATATCTACACATACAATATAGATGTTTTCATATAAATGCTTTCATCTTAACAGAAAT	1500
Db	1316	ATATCTACACATACAATATAGATGTTTTCATATAAATGCTTTCATCTTAACAGAAAT	1375
Qy	1501	GTCTTTTGTGATTTGTTTTAGAAAACCTGAGAGTTTTAAATTCATAAATACTTTGATCAAAAA	1560
Db	1376	GTCTTTTGTGATTTGTTTTAGAAAACCTGAGAGTTTTAAATTCATAAATACTTTGATCAAAAA	1435
Qy	1561	TTGTGGGAACAATCCAGCAATAAATGTATGTGATTTGTTTTATGTACATAAGGAGTCTTA	1620
Db	1436	TTGTGGGAACAATCCAGCAATAAATGTATGTGATTTGTTTTATGTACATAAGGAGTCTTA	1495
Qy	1621	AGCTTTGTGGCTTTGAAGTCTTTTGTACTTATAGTCCCATGTTTAAAAATTACTTATATC	1680
Db	1496	AGCTTTGTGGCTTTGAAGTCTTTTGTACTTATAGTCCCATGTTTAAAAATTACTTATATC	1555
Qy	1681	TAAAGCAATTTATGTTTTTCAATTTCAATTTACATGATGCTAAATTTATGGCAATTTATAACAA	1740
Db	1556	TAAAGCAATTTATGTTTTTCAATTTCAATTTACATGATGCTAAATTTATGGCAATTTATAACAA	1615
Qy	1741	TATTAAGAATTCGAATAGAAATATGTGAATTTGTTCACATACATAGAAATGAAAGTTCA	1800
Db	1616	TATTAAGAATTCGAATAGAAATATGTGAATTTGTTCACATACATAGAAATGAAAGTTCA	1675
Qy	1801	TTTCGTAAAGCAAGATGCTGGGTGAAAGAGTGCTTTTGATTTGAAAGATCACTAGATTAGT	1860
Db	1676	TTTCGTAAAGCAAGATGCTGGGTGAAAGAGTGCTTTTGATTTGAAAGATCACTAGATTAGT	1735
Qy	1861	AGAGGGCAAGACTCTAGTCCCTAATCTACCCCTTAATAGCCATGTGGTCACGTGTAAAGTC	1920
Db	1736	AGAGGGCAAGACTCTAGTCCCTAATCTACCCCTTAATAGCCATGTGGTCACGTGTAAAGTC	1795
Qy	1921	AGTGAACCCATCTATCTCCTCATACTTTTTTTCATCTCTAAAATAGGGTATAATTTAA	1980
Db	1796	AGTGAACCCATCTATCTCCTCATACTTTTTTTCATCTCTAAAATAGGGTATAATTTAA	1855
Qy	1981	GCCTTTCATTTTTTTTTTTTTTTTTTTTTTTTGAGATAGATTTTGCTCTTGTCACCCAGTTGGAGTC	2040
Db	1856	GCCTTTCATTTTTTTTTTTTTTTTTTTTTTTTGAGATAGATTTTGCTCTTGTCACCCAGTTGGAGTC	1915
Qy	2041	CAATGGCAGCATCTACGCTCACTGCAACCTGCTTCTCCTCGGTTCAAGTGATTTCCCTG	2100
Db	1916	CAATGGCAGCATCTACGCTCACTGCAACCTGCTTCTCCTCGGTTCAAGTGATTTCCCTG	1975
Qy	2101	CTTCAGCCTCCCAAGTAGCCGGGATTACAGGTGCCGCCACCCACATCTGGCTAAATTTTTT	2160
Db	1976	CTTCAGCCTCCCAAGTAGCCGGGATTACAGGTGCCGCCACCCACATCTGGCTAAATTTTTT	2035
Qy	2161	GTATTTTACCATGTTGGCCAGGCTGGTCTCGAACCCCTACCTCAGGTGATCCCTCGCCT	2220
Db	2036	GTATTTTACCATGTTGGCCAGGCTGGTCTCGAACCCCTACCTCAGGTGATCCCTCGCCT	2095
Qy	2221	CGGCTCTCAAAGTGTCTGGGATTACAGGTGTGAGCCACCGCCCAATATCAGTT	2280
Db	2096	CGGCTCTCAAAGTGTCTGGGATTACAGGTGTGAGCCACCGCCCAATATCAGTT	2155
Qy	2281	TTTTCTTTTTTAAACACRAGGCTAACACAAATCAAAATACTAGTAGGGGAGAAAAAAAAT	2340

Db	2156	 TTTTCTTTTAAACAAGGCTAAACAATCAAAATACTAGCTAGGGGAGAAAAAANAAT	2215
Qy	2341	AAGGCACGTGTTATGTATACAGGCTCTGTTGCCAATCACTGGCGACAGAAATAAACAGTA	2400
Db	2216	AAGGCACGTGTTATGTATACAGGCTCTGTTGCCAATCACTGGCGACAGAAATAAACAGTA	2275
Qy	2401	AGAAATCAATCCTTTTCATATATCCTTTCTTGCGAAGATACATAAAATCCCACAAATGCGTAT	2460
Db	2276	AGAAATCAATCCTTTTCATATATCCTTTCTTGCGAAGATACATAAAATCCCACAAATGCGTAT	2335
Qy	2461	CTTCCCTTTTATGATATTTGGAGAAATTGTAAGTGACAGATATTTTCTTGGGTGTA	2520
Db	2336	CTTCCCTTTTATGATATTTGGAGAAATTGTAAGTGACAGATATTTTCTTGGGTGTA	2395
Qy	2521	TAGACCACAAGGACGTGTTTTCATGATGCTTTGTCATATAAAATATACCTTAGTTTTFACT	2580
Db	2396	TAGACCACAAGGACGTGTTTTCATGATGCTTTGTCATATAAAATATACCTTAGTTTTFACT	2455
Qy	2581	TTGTATGTTACATGTTAGATTTAGAGTATGAAAATTAGTAGGGAGGATTTATAACAAAGA	2640
Db	2456	TTGTATGTTACATGTTAGATTTAGAGTATGAAAATTAGTAGGGAGGATTTATAACAAAGA	2515
Qy	2641	ACAGGGCAAGGAGGTAGAAATTAACCTCTTAATACCTGTGCACAAGTAGGCTTTTCA	2700
Db	2516	ACAGGGCAAGGAGGTAGAAATTAACCTCTTAATACCTGTGCACAAGTAGGCTTTTCA	2575
Qy	2701	GAACACTCAACCCCTACATAAACTGGATAGTTAGAAAAGCACACTCCCAAGGAAGCGG	2760
Db	2576	GAACACTCAACCCCTACATAAACTGGATAGTTAGAAAAGCACACTCCCAAGGAAGCGG	2635
Qy	2761	TTATGTTTTGCAGTTTGAATCAGAAGAAATAGAGCTATAGCAATCTTCATTCTATAGTAAC	2820
Db	2636	TTATGTTTTGCAGTTTGAATCAGAAGAAATAGAGCTATAGCAATCTTCATTCTATAGTAAC	2695
Qy	2821	ATTAAAGAGCCTGGTTTATATATAGCAGTCATTAAGATTTTAAAAATTTACATCTTGCCG	2880
Db	2696	ATTAAAGAGCCTGGTTTATATATAGCAGTCATTAAGATTTTAAAAATTTACATCTTGCCG	2755
Qy	2881	TTCTCTCTTACTCACAGATTTTCGAGAGGTAACTGAATGATCCACAGGTCAGAAATCACTG	2940
Db	2756	TTCTCTCTTACTCACAGATTTTCGAGAGGTAACTGAATGATCCACAGGTCAGAAATCACTG	2815
Qy	2941	CTTTTTATAATGCGATTAAATTCGATGAACAAAGTTTCCACAAATAAACAGTAATAAAAA	3000
Db	2816	CTTTTTATAATGCGATTAAATTCGATGAACAAAGTTTCCACAAATAAACAGTAATAAAAA	2875
Qy	3001	GAACAATGTAATAGCACTTAAATAAGCCAGGGCTGGACGACGTGTGTTACATGCTTTCAA	3060
Db	2876	GAACAATGTAATAGCACTTAAATAAGCCAGGGCTGGACGACGTGTGTTACATGCTTTCAA	2935
Qy	3061	TCCATGAACGTGTAATGCTACTAGTATCTCTATTGGACATCTGAGGAACCAATGGA	3120
Db	2936	TCCATGAACGTGTAATGCTACTAGTATCTCTATTGGACATCTGAGGAACCAATGGA	2995
Qy	3121	GTTTGATAACAGTAGAGTTAAAAATTAATCATATATATATATATGCTCAATCTCACAG	3180
Db	2996	GTTTGATAACAGTAGAGTTAAAAATTAATCATATATATATATATGCTCAATCTCACAG	3055
Qy	3181	ACATCTCTGTACCAAAAGCTATCATATCTAGATATGCGGCATPAAGATGACCTTGGGC	3240
Db	3056	ACATCTCTGTACCAAAAGCTATCATATCTAGATATGCGGCATPAAGATGACCTTGGGC	3115
Qy	3241	ACACTAGATTCTTTGAGAGAAATCTGGCAGAGAAAACAAATATTTATTCCTACAAATAA	3300
Db	3116	ACACTAGATTCTTTGAGAGAAATCTGGCAGAGAAAACAAATATTTATTCCTACAAATAA	3175
Qy	3301	ACCCAGCATTTTACAGGTTTTATTTTTTAACATGAAGTATTGTTATCTGTATCTTTCATA	3360
Db	3176	ACCCAGCATTTTACAGGTTTTATTTTTTAACATGAAGTATTGTTATCTGTATCTTTCATA	3235
Qy	3361	TAAGTGTGCCGGAAATTTATTTCTTCTGTTGGGTCTTGCTCTGCTGACTCCAAAGAT	3420



Db	3236	TAAGTGTGCCGGAAATTAATTTCTTCTGGTGGGTTCTTGGTCTCGCTGACTCCAGAATG	3295	Qy	4501	ATCAACTGAATGTAAATAATTTCCAGGCATCCCTGTTATTTATTTATTTGTCT	4560
Qy	3421	AAACCGCAGACCCCTTGAGGTGAGTGTACAGTTCCTTAAAGATGCTGTCTTCAGAGTTGT	3480	Db	4376	ATCAACTGAATGTAAATAATTTCCAGGCATCCCTGTTATTTATTTATTTGTCT	4435
Db	3296	AAACCGCAGACCCCTTGAGGTGAGTGTACAGTTCCTTAAAGATGCTGTCTTCAGAGTTGT	3355	Qy	4561	CTCTTTCTGCTGCTTACTTCAAAAAGTCATATGGCATGGTGAAGTGTCTGCTGCC	4620
Qy	3481	TCCTTTCAGATGTTTCAGATGTTCCGGAGTTTCTCCCTTATGTTGAGTTCGTGGTCTCGCT	3540	Db	4436	CTCTTTCTGCTGCTTACTTCAAAAAGTCATATGGCATGGTGAAGTGTCTGCTGCC	4495
Db	3356	TCCTTTCAGATGTTTCAGATGTTCCGGAGTTTCTCCCTTATGTTGAGTTCGTGGTCTCGCT	3415	Qy	4621	AAAGAGTTTGTCAATATAAGATTCCTTTCTTTGTAACAATTCATCTACCTTGGGGCTTCATTT	4680
Qy	3541	GACTTCAACAAATGAAGCCGACAGCTTTGAGTGTGAGTGTGTGACAGTTCCTTAAAGGAGT	3600	Db	4496	AAAGAGTTTGTCAATATAAGATTCCTTTCTTTGTAACAATTCATCTACCTTGGGGCTTCATTT	4555
Db	3416	GACTTCAACAAATGAAGCCGACAGCTTTGAGTGTGAGTGTGTGACAGTTCCTTAAAGGAGT	3475	Qy	4681	ATATCAAAAAGGAGTACTGTAAACCTGTCAAAAAAGCTACTCTGTGCAATATATATATGT	4740
Qy	3601	GGGTCCAGAGTGTGTTCTCTCCCGTAGTGTCTGCTGCTGCTGATGTGACGAATGAA	3660	Db	4556	ATATCAAAAAGGAGTACTGTAAACCTGTCAAAAAAGCTACTCTGTGCAATATATATATGT	4615
Db	3476	GGGTCCAGAGTGTGTTCTCTCCCGTAGTGTCTGCTGCTGCTGATGTGACGAATGAA	3535	Qy	4741	GATGGTTTACCTGTCAGTAAGGTGGTGGCAATAAATAAATCAACAGTAATGAAACCG	4800
Qy	3661	GCTGCAGACCTCGCGGTAAAGTGTACAGCTCATATAAGGTAGTGCATAACCCCAACAGTGA	3720	Db	4616	GATGGTTTACCTGTCAGTAAGGTGGTGGCAATAAATAAATCAACAGTAATGAAACCG	4675
Db	3536	GCTGCAGACCTCGCGGTAAAGTGTACAGCTCATATAAGGTAGTGCATAACCCCAACAGTGA	3595	Qy	4801	AGCAGAACTGTCAGAGAAATGGTTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAA	4860
Qy	3721	GCAGTAGCAAGATTTATATGAGAGCAAAAGAAAGCTTCCCAACCATAGAACGGA	3780	Db	4676	AGCAGAACTGTCAGAGAAATGGTTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAA	4735
Db	3596	GCAGTAGCAAGATTTATATGAGAGCAAAAGAAAGCTTCCCAACCATAGAACGGA	3655	Qy	4861	TAATCATATAAATAGCCACTCAATCCAAAACATCACTGGCGCACTTGTACATATATAA	4920
Qy	3781	CCAGAATTGGTGTGCTGCTGTGTTAGCCAGCTTTTATCCCTTATTTGCGCCACACCCA	3840	Db	4736	TAATCATATAAATAGCCACTCAATCCAAAACATCACTGGCGCACTTGTACATATATAA	4795
Db	3656	CCAGAATTGGTGTGCTGCTGTGTTAGCCAGCTTTTATCCCTTATTTGCGCCACACCCA	3715	Qy	4921	TCAGTGGAGATGTGATTGAAGCACAAGGCTTAAAGTGAATGTCTAGAGAGCTAAATTGATTC	4980
Qy	3841	CATCCCTGCTGATTGGCCATTTTACAGAATGCTGATTGGTCCATTTTATAGCGTGTGAT	3900	Db	4796	TCAGTGGAGATGTGATTGAAGCACAAGGCTTAAAGTGAATGTCTAGAGAGCTAAATTGATTC	4855
Db	3716	CATCCCTGCTGATTGGCCATTTTACAGAATGCTGATTGGTCCATTTTATAGCGTGTGAT	3775	Qy	4981	ATTTTATGAGAAATTTTACTTATTTTAAATGTCATCCCTGACCATCTTGAACCTTTACTT	5040
Qy	3901	TGGTGGCTTTTACAGAGTCTGATGTTGGTGCATTTTACAATCCCTTTAGCTAGACACAGT	3960	Db	4856	ATTTTATGAGAAATTTTACTTATTTTAAATGTCATCCCTGACCATCTTGAACCTTTACTT	4915
Db	3776	TGGTGGCTTTTACAGAGTCTGATGTTGGTGCATTTTACAATCCCTTTAGCTAGACACAGT	3835	Qy	5041	GAAGATTTATTTTATTTTAAATCACTGTTTATAGATTTAGGTATTTCTGGTCTTTGTT	5100
Qy	3961	GCTGATTGGTGCCTTTTATAATCCTTTAGCTAGACAAAAAGTCTACAAGTCCCAACCCA	4020	Db	4916	GAAGATTTATTTTATTTTAAATCACTGTTTATAGATTTAGGTATTTCTGGTCTTTGTT	4975
Db	3836	GCTGATTGGTGCCTTTTATAATCCTTTAGCTAGACAAAAAGTCTACAAGTCCCAACCCA	3895	Qy	5101	TTTCTTTTATCATGATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5160
Qy	4021	ACCCAGAAGTCCGCTGGCTTACCTCTCTGTAAGGAAATTTAGGTTTCAAAACAAGTTTCAA	4080	Db	4976	TTTCTTTTATCATGATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5035
Db	3896	ACCCAGAAGTCCGCTGGCTTACCTCTCTGTAAGGAAATTTAGGTTTCAAAACAAGTTTCAA	3955	Qy	5161	TGAGAAGAAATGATTAATAATCCATTTATTTTACCTTACCTTACCTTACCTTACCTTACCTT	5217
Qy	4081	AGTGCTAAAACACTACAGTTTCTCATTTCTGCAACTGGATTTCCACTCATGTTTGAATC	4140	Db	5036	TGAGAAGAAATGATTAATAATCCATTTATTTTACCTTACCTTACCTTACCTTACCTTACCTT	5092
Db	3956	AGTGCTAAAACACTACAGTTTCTCATTTCTGCAACTGGATTTCCACTCATGTTTGAATC	4015	RESULT 3			
Qy	4141	CCAGGCTCTAAGACTTAACCTTGCATTTCTGTGCTTATGTTCTGCAATTTACACAAAG	4200	US-09-996-956-5			
Db	4016	CCAGGCTCTAAGACTTAACCTTGCATTTCTGTGCTTATGTTCTGCAATTTACACAAAG	4075	; Sequence 5, Application US/09996956			
Qy	4201	CTACTATCTGCATCTCTGTTGTTAACTTCAGACTTAAACTTCTTTTGTGATTCACAATG	4260	; Patent No. US20020155463A1			
Db	4076	CTACTATCTGCATCTCTGTTGTTAACTTCAGACTTAAACTTCTTTTGTGATTCACAATG	4135	; GENERAL INFORMATION:			
Qy	4261	ACCACACACTTTTGGTGGAGTTTGTCTATCGGTTTATGTTACTGGTTAATAGAGACT	4320	; APPLICANT: Origene Technologies, Inc			
Db	4136	ACCACACACTTTTGGTGGAGTTTGTCTATCGGTTTATGTTACTGGTTAATAGAGACT	4195	; TITLE OF INVENTION: Prostate Polynucleotides and Uses			
Qy	4321	TCCTTCCAGAAATTTGAGTAGATGAAGAGGAAGTAGCACATTCCTTAAATAATGTACCATGC	4380	; FILE REFERENCE: 9U 301 R1			
Db	4196	TCCTTCCAGAAATTTGAGTAGATGAAGAGGAAGTAGCACATTCCTTAAATAATGTACCATGC	4255	; CURRENT APPLICATION NUMBER: US/09/996,956			
Qy	4381	CTTTCAAGTCACAAGCATCCCTATCATATGGCTGTCAAGGGTGGCTCAGAAATAGGTAGAG	4440	; CURRENT FILING DATE: 2001-11-30			
Db	4256	CTTTCAAGTCACAAGCATCCCTATCATATGGCTGTCAAGGGTGGCTCAGAAATAGGTAGAG	4315	; PRIOR APPLICATION NUMBER: 60/250,354			
Qy	4441	TTAAGAAATTTAAAGTAAATTTGGTGTGAAGCATGAAGGCTTCAATCTTAAAGCTTATATTAC	4500	; PRIOR FILING DATE: 2001-12-01			
Db	4316	TTAAGAAATTTAAAGTAAATTTGGTGTGAAGCATGAAGGCTTCAATCTTAAAGCTTATATTAC	4375	; NUMBER OF SEQ ID NOS: 16			
				; SOFTWARE: PatentIn version 3.1			
				; SEQ ID NO 5			
				; LENGTH: 32463			
				; TYPE: DNA			
				; ORGANISM: Homo sapiens			
				US-09-996-956-5			
				Query Match 91.2%; Score 4758.4; DB 10; Length 32463;			
				Best Local Similarity 99.8%; Pred. No. 0;			
				Matches 4765; Conservative 0; Mismatches 11; Indels 0; Gaps 0;			

Qy	423	AACTAACTGGGATTCACAGAGGCTGGTGCAATTCACACTTCATGATATCAGCCAAAGTG	482
Db	27688	AAATATTTCTCACCCAGCAGGCTGGTGCAATTCACACTTCATGATATCAGCCAAAGTG	27747
Qy	483	GAACTAAAACAGCTCCTGGAAGAGGACTATCACATCATCAGGTGGGAGTCTCCAGGA	542
Db	27748	GAACTAAAACAGCTCCTGGAAGAGGACTATCACATCATCAGGTGGGAGTCTCCAGGA	27807
Qy	543	CAGCGGACCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAAT	602
Db	27808	CAGCGGACCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAAT	27867
Qy	603	TCTCTGTCCTGTAAAGCATTTCATATTTACAAGACACAGGCCCTACTCTTAGGCCAGCA	662
Db	27868	TCTCTGTCCTGTAAAGCATTTCATATTTACAAGACACAGGCCCTACTCTTAGGCCAGCA	27927
Qy	663	AAAAGTGGCAACAGGCAACGAGGGAAGAGATCATGAGGCATTTTCAGAGTGCACHTGT	722
Db	27928	AAAAGTGGCAACAGGCAACGAGGGAAGAGATCATGAGGCATTTTCAGAGTGCACHTGT	27987
Qy	723	CTTTTCATATATTTCTCAATCCGTATGTTTGGTTTTATTTGGCCAAAGCATAAACAATCT	782
Db	27988	CTTTTCATATATTTCTCAATCCGTATGTTTGGTTTTATTTGGCCAAAGCATAAACAATCT	28047
Qy	783	GCTCAAGAAAAAAAATCTGGAGAAAAACAAAGTGCCCTTGCCAAATGTTATGTTCTCTTTT	842
Db	28048	GCTCAAGAAAAAAAATCTGGAGAAAAACAAAGTGCCCTTGCCAAATGTTATGTTCTCTTTT	28107
Qy	843	TGACAAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGATCAGGTCAATTCATTTAC	902
Db	28108	TGACAAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGATCAGGTCAATTCATTTAC	28167
Qy	903	GTTGTGTCAAAATATGATTTTAAAGATPACAACTTTGCAGAGAGCATGCTTCTCTAAGGT	962
Db	28168	GTTGTGTCAAAATATGATTTTAAAGATPACAACTTTGCAGAGAGCATGCTTCTCTAAGGT	28227
Qy	963	AGGCAGCTGGAGACTTAAGGTAAAGCATTTCAAGATCAGTTAATCAAGAAAGTGCT	1022
Db	28228	AGGCAGCTGGAGACTTAAGGTAAAGCATTTCAAGATCAGTTAATCAAGAAAGTGCT	28287
Qy	1023	CTTTGCAATCTGAAATGCCCCTTTGTGCAAAATATGTTTAAATTTACACTTA	1082
Db	28288	CTTTGCAATCTGAAATGCCCCTTTGTGCAAAATATGTTTAAATTTACACTTA	28347
Qy	1083	ATGGAACAACCTTTAACTTCAGATGAACAAACCCACAAAGCAAAAATCAAAAGCCC	1142
Db	28348	ATGGAACAACCTTTAACTTCAGATGAACAAACCCACAAAGCAAAAATCAAAAGCCC	28407
Qy	1143	TACCTATGATTTTCATATTTCTGTGTAACGTGATTAAGGATTCCTGCTGCTTTGGGC	1202
Db	28408	TACCTATGATTTTCATATTTCTGTGTAACGTGATTAAGGATTCCTGCTGCTTTGGGC	28467
Qy	1203	ATAAATGATAATGGAAATTTCCAGGTATGTTTAAATGAGGGCCCATCTACAAATCT	1262
Db	28468	ATAAATGATAATGGAAATTTCCAGGTATGTTTAAATGAGGGCCCATCTACAAATCT	28527
Qy	1263	TAGCAATCTTTGGATAATTTCAAATTCAGCTGGACATGTCATATGTTTTTATATA	1322
Db	28528	TAGCAATCTTTGGATAATTTCAAATTCAGCTGGACATGTCATATGTTTTTATATA	28587
Qy	1323	CATCTTTGCTAGAAATTTCAAAATTTTAAAGTATGTGAATTTAGTTAATAGCTGCTGATC	1382
Db	28588	CATCTTTGCTAGAAATTTCAAAATTTTAAAGTATGTGAATTTAGTTAATAGCTGCTGATC	28647
Qy	1383	AATTCAAAAACATTAATTTCTTAAATTTTAGACTATGAAGGTCAATAAATCAACAATAT	1442
Db	28648	AATTCAAAAACATTAATTTCTTAAATTTTAGACTATGAAGGTCAATAAATCAACAATAT	28707
Qy	1443	ATCTACACATACAATTAATAGATGTTTTTCATTAATAATGCTTCACTTAACAGAAATGT	1502
Db	28708	ATCTACACATACAATTAATAGATGTTTTTCATTAATAATGCTTCACTTAACAGAAATGT	28767

Qy	1503	CTTTGTGATGTGTTTTAGAAAACGTGAGAGTTTTAAATTCATTAATTAATCTGATCAAAAATTT	1566
Db	28768	CTTTGTGATGTGTTTTAGAAAACGTGAGAGTTTTAAATTCATTAATTAATCTGATCAAAAATTT	28822
Qy	1563	GTGGGAACAATCCAGCATTAATTTGATGTGATTTGTTTTATGTACATAAGGAGTCTTAAG	1622
Db	28828	GTGGGAACAATCCAGCATTAATTTGATGTGATTTGTTTTATGTACATAAGGAGTCTTAAG	28888
Qy	1623	CTTGGTGCCCTGAAGTCTTTTTGTACTTAGTCCCATGTTTAAAAATTAATCTTATAATCTA	1682
Db	28888	CTTGGTGCCCTGAAGTCTTTTTGTACTTAGTCCCATGTTTAAAAATTAATCTTATAATCTA	28948
Qy	1683	AAGCATTTATGTTTTTCAATTCATTAATTTACATGATGCTAATTTATGGCAAATTAACAAATA	1742
Db	28948	AAGCATTTATGTTTTTCAATTTCAATTTACATGATGCTAATTTATGGCAAATTAACAAATA	29008
Qy	1743	TTAAAGATTCGAAATAGAAATATGTGAATTTTCACATACATAGAAATGAAAGTTTCATTT	1802
Db	29008	TTAAAGATTCGAAATAGAAATATGTGAATTTTCACATACATAGAAATGAAAGTTTCATTT	29066
Qy	1803	TCGTAAAGCAAGATGCTGGGTGAAGAGTGCTTTTCATTTGAAGATCACTAGATTAGTAG	1862
Db	29068	TCGTAAAGCAAGATGCTGGGTGAAGAGTGCTTTTCATTTGAAGATCACTAGATTAGTAG	29127
Qy	1863	AGGCAAGACTTCTAGTCCCTAATCTACCCCTTAATAGCCATGTGGTCAGGTGTAAGTCAG	1922
Db	29128	AGGCAAGACTTCTAGTCCCTAATCTACCCCTTAATAGCCATGTGGTCAGGTGTAAGTCAG	29188
Qy	1923	TGAACCCATCTCATTTCTCTCATACACTTTTTTCATCTCTAAAATGAGGGTATAATTTAAGC	1982
Db	29188	TGAACCCATCTCATTTCTCTCATACACTTTTTTCATCTCTAAAATGAGGGTATAATTTAAGC	29248
Qy	1983	TCCTTCATTTTTTTTTTTTTTTGAGATAGAGTTTTGCTCTTGTCACCCACAGTTGGAGTGCA	2042
Db	29248	TCCTTCATTTTTTTTTTTTTTTGAGATAGAGTTTTGCTCTTGTCACCCACAGTTGGAGTGCA	29307
Qy	2043	ATGGCAGATCTCAGTCTCACTGCAACCCCTCTGCTTCTCTCGGTTCAAGTGATCTCTCTGCT	2102
Db	29308	ATGGCAGATCTCAGTCTCACTGCAACCCCTCTGCTTCTCTCGGTTCAAGTGATCTCTCTGCT	29366
Qy	2103	TCAGCCTCCCAAGTAGCCGGATTTACAGGTGCCGCCACACATCTGGCTAAATTTTTTCT	2162
Db	29368	TCAGCCTCCCAAGTAGCCGGATTTACAGGTGCCGCCACACATCTGGCTAAATTTTTTCT	29427
Qy	2163	ATTTTCACCATGTTGGCCAGGCTGGTCTCGAACCCCTCACTCAGGTGATCCCTCGCCTCG	2222
Db	29428	ATTTTCACCATGTTGGCCAGGCTGGTCTCGAACCCCTCACTCAGGTGATCCCTCGCCTCG	29488
Qy	2223	GCCTCTCAAGTCTGGGATTTACAGTGTGAGCCACCGCCAGCCCAATATCAGTTTTT	2282
Db	29488	GCCTCTCAAGTCTGGGATTTACAGTGTGAGCCACCGCCAGCCCAATATCAGTTTTT	29547
Qy	2283	TCTTTTTTTAAACAAGCTAACACAATCAAAAATTAAGTTAGGGGAGAAAAAATAA	2342
Db	29548	TCTTTTTTTAAACAAGCTAACACAATCAAAAATTAAGTTAGGGGAGAAAAAATAA	29607
Qy	2343	GGCAGCTTTTATGTGTAAACAGGCTCTTTTGTCAATCACTTGGCCAGACAAATAACAGTAG	2402
Db	29608	GGCAGCTTTTATGTGTAAACAGGCTCTTTTGTCAATCACTTGGCCAGACAAATAACAGTAG	29667
Qy	2403	AATCAATCCCTTTTTCATATCTCTTTGCAAGATACATAAAATCCCAAAATGGCTATCT	2462
Db	29668	AATCAATCCCTTTTTCATATCTCTTTGCAAGATACATAAAATCCCAAAATGGCTATCT	29727
Qy	2463	TCCTTTTTTATGATTTTGGAGAAATGTAGCTAAGTGACAGATATTTTTGCTTGGGTGTATA	2522
Db	29728	TCCTTTTTTATGATTTTGGAGAAATGTAGCTAAGTGACAGATATTTTTGCTTGGGTGTATA	29787
Qy	2523	GACCAACAAGGACTGTGTTTTGATGATGGTTTGCATAAAATATACCTTAGTTTTTACTTTT	2582
Db	29788	GACCAACAAGGACTGTGTTTTGATGATGGTTTGCATAAAATATACCTTAGTTTTTACTTTT	29847
Qy	2583	GTATGTTTACATGTTTAGATTTTAGAGTATGAAAAATTAGTAGGGAGGATTAATTAACAAAGAAC	2642

|||||  
29848 Db GTATGTTACATGTTAGATTTAGAGTATCAAAATTTAGTAGGGAGGATTATTAAACAAGAAC 29907  
|||||  
2643 QY AGGCGAAGAGAGTAGAATTAACACTCTTCTAAATACCTGTGCACAAAGTAGGCTTTTCAGA 2702  
|||||  
29908 Db AGGCGAAGAGAGTAGAATTAACACTCTTCTAAATACCTGTGTGCACAAAGTAGGCTTTTCAGA 29967  
|||||  
2703 QY AACTCTACAAACCCCTACATAAACTGGATAGTTAGAAAAGCACACACCTCCCAAGGAAGCGGTT 2762  
|||||  
29968 Db AACTCTACAAACCCCTACATAAACTGGATAGTTAGAAAAGCACACACCTCCCAAGGAAGCGGTT 30027  
|||||  
2763 QY ATGTTTTGCAGTTTGAATTCAGAAGAATAGAGCTATAGCAATCTTCTCATTTCTATAGTAACAT 2822  
|||||  
30028 Db ATGTTTTGCAGTTTGAATTCAGAAGAATAGAGCTATAGCAATCTTCTCATTTCTATAGTAACAT 30087  
|||||  
2823 QY TAAAGAGCCGGTTTATATATATAGCAGTCATTAAGATTTAAATAATTTACATCTTGCCTTT 2882  
|||||  
30088 Db TAAAGAGCCGGTTTATATATATAGCAGTCATTAAGATTTAAATAATTTACATCTTGCCTTT 30147  
|||||  
2883 QY CTTCTTACTCACAGATTTTCGAGAGGTAATGTAATGATCCACGAGGTGAGAATCACCTGCC 2942  
|||||  
30148 Db CTTCTTACTCACAGATTTTCGAGAGGTAATGTAATGATCCACGAGGTGAGAATCACCTGCC 30207  
|||||  
2943 QY TTTTATAATCGGATTAATTCATGAACAAAGTTTCCAAACAAATAACAGTAATAAAAGA 3002  
|||||  
30208 Db TTTTATAATCGGATTAATTCATGAACAAAGTTTCCAAACAAATAACAGTAATAAAAGA 30267  
|||||  
3003 QY AACATGATTTAGCACTTAATAAGCCAGGGGTGCACGACGTGTTCATGCTTTCATC 3062  
|||||  
30268 Db AACATGATTTAGCACTTAATAAGCCAGGGGTGCACGACGTGTTCATGCTTTCATC 30327  
|||||  
3063 QY CATGAACCTGGTAAACTGCTACTAGTATCTCTATTGGACATGTGAGGAACCAATGCGAGT 3122  
|||||  
30328 Db CATGAACCTGGTAAACTGCTACTAGTATCTCTATTGGACATGTGAGGAACCAATGCGAGT 30387  
|||||  
3123 QY TGATAAACAGTAGAGTTTAAATAATTTACTCTTCATATATATATATGCTTCAATCTCACAGAC 3182  
|||||  
30388 Db TGATAAACAGTAGAGTTTAAATAATTTACTCTTCATATATATATATGCTTCAATCTCACAGAC 30447  
|||||  
3183 QY ATCTCTGCTACCAAAAGCTATCATATCTAGATATGCGGCATAGGATGACCTTGGGGCAC 3242  
|||||  
30448 Db ATCTCTGCTACCAAAAGCTATCATATCTAGATATGCGGCATAGGATGACCTTGGGGCAC 30507  
|||||  
3243 QY ACTAGAAATCTTTGAGAGAAATCTGCGAGAGAAACAAATATTTATTCCTACATAAAG 3302  
|||||  
30508 Db ACTAGAAATCTTTGAGAGAAATCTGCGAGAGAAACAAATATTTATTCCTACATAAAG 30567  
|||||  
3303 QY CCAGCATTTTACAGGTTTTATTTTAACTATGAAGTATGTTATCTGTATCTTTTCATATA 3362  
|||||  
30568 Db CCAGCATTTTACAGGTTTTATTTTAACTATGAAGTATGTTATCTGTATCTTTTCATATA 30627  
|||||  
3363 QY AGTGTGCCCGGAATTTATTTCTTCTGTTGGGTTCTTGGTCTCGCTGACTCCAAGATGAA 3422  
|||||  
30628 Db AGTGTGCCCGGAATTTATTTCTTCTGTTGGGTTCTTGGTCTCGCTGACTCCAAGATGAA 30687  
|||||  
3423 QY ACCGCGAGACCTTTGAGGTGAGTGTCAAGTTCCTTAAAGATGTTGTGTTCAGAGTTTGTTC 3482  
|||||  
30688 Db ACCGCGAGACCTTTGAGGTGAGTGTCAAGTTCCTTAAAGATGTTGTGTTCAGAGTTTGTTC 30747  
|||||  
3483 QY CTTTCAGATGTTTCAGATGTGTCGAGGTTTCTCCCTTATGGTGAAGTTCGTGCTCGCTGA 3542  
|||||  
30748 Db CTTTCAGATGTTTCAGATGTGTCGAGGTTTCTCCCTTATGGTGAAGTTCGTGCTCGCTGA 30807  
|||||  
3543 QY CTTTCAACAATGAAGCCGAGACCTTTCAGTGAAGTGTGTCACAGTTCCTTAAAGCGAGTGC 3602  
|||||  
30808 Db CTTTCAACAATGAAGCCGAGACCTTTCAGTGAAGTGTGTCACAGTTCCTTAAAGCGAGTGC 30867  
|||||  
3603 QY GTCCAGAGTTGTTTGTCTCCCGTAGGTTCTCGCTGCTCGCTGATGTGAGGAATGAAGC 3662  
|||||  
30868 Db GTCCAGAGTTGTTTGTCTCCCGTAGGTTCTCGCTGCTCGCTGATGTGAGGAATGAAGC 30927  
|||||  
3663 QY TGCAGACCCCTCGCGGTAAAGTTTACAGTTCATAAAGGTAGTGAACCCCAACAGTGAGC 3722  
|||||

30928 Db TGCAGACCCCTCGCGTAAAGTGTACAGCTCATAAAGGTAGTGCAAAACCAACAGTAGTGAGC 30987  
|||||  
3723 QY AGTAGCAAGATTTTATTATGAGAGCAAAAGAACAAAGCTTCCCCACCATAGAAACGGACC 3782  
|||||  
30988 Db AGTAGCAAGATTTTATTATGAGAGCAAAAGAACAAAGCTTCCCCACCATAGAAACGGACC 31047  
|||||  
3783 QY AGAATTGGTGTCTGCTGCTGTGGTAGCCAGCTTTTATTTCCCTTATTTTGGCCACACCCACA 3842  
|||||  
31048 Db AGAATTGGTGTCTGCTGCTGTGGTAGCCAGCTTTTATTTCCCTTATTTTGGCCACACCCACA 31107  
|||||  
3843 QY TCCTGTGATTTGGCCCCATTTTACAGAATGCTGATTTGGTTCCTATTTTATAGCGTGCTGATG 3902  
|||||  
31108 Db TCCTGTGATTTGGCCCCATTTTACAGAATGCTGATTTGGTTCCTATTTTATAGCGTGCTGATG 31167  
|||||  
3903 QY GTGGGTTTTTACAGAGTGCTGATTTGGTGCAATTTTACAATCTTTTACCTTAGCTAGACACAGATGC 3962  
|||||  
31168 Db GTGGGTTTTTACAGAGTGCTGATTTGGTGCAATTTTACAATCTTTTACCTTAGCTAGACACAGATGC 31227  
|||||  
3963 QY TGATTGGTGTCTTTTATAATCTCTTTAGCTAGACACAAAGTTTCTACAAGTCCCCACCCCAAC 4022  
|||||  
31228 Db TGATTGGTGTCTTTTATAATCTCTTTAGCTAGACACAAAGTTTCTACAAGTCCCCACCCCAAC 31287  
|||||  
4023 QY CCAGAAGCTCGCTGGCTTCCACCTCTCGTAAGAAATTTAGAGTTTCAACAAGTTTCAAG 4082  
|||||  
31288 Db CCAGAAGCTCGCTGGCTTCCACCTCTCGTAAGAAATTTAGAGTTTCAACAAGTTTCAAG 31347  
|||||  
4083 QY TGCTAAAACCTACAGTTTCTCATTTCTCTGCAACTGGGATTTCCACCTCATGTGTTTGAATCCC 4142  
|||||  
31348 Db TGCTAAAACCTACAGTTTCTCATTTCTCTGCAACTGGGATTTCCACCTCATGTGTTTGAATCCC 31407  
|||||  
4143 QY AGGCTCTAAGACCTTAACTTGCCATTCTGACTTCTATGTTTCTCTGCAATTTTACACAAAGCT 4202  
|||||  
31408 Db AGGCTCTAAGACCTTAACTTGCCATTCTGACTTCTGACTTCTGCTGCAATTTTACACAAAGCT 31467  
|||||  
4203 QY ACTACTGTGCACATCTCTGCTGTTAACTTCCAGACTAAACCTCTTTTGTGATTCACAAATGAC 4262  
|||||  
31468 Db ACTACTGTGCACATCTCTGCTGTTAACTTCCAGACTAAACCTCTTTTGTGATTCACAAATGAC 31527  
|||||  
4263 QY CACACACTTTTGGTGTAGGTTTGTCTATCGGTTTATTTGTTACTGGTTTAAAGAGAGCTTC 4322  
|||||  
31528 Db CACACACTTTTGGTGTAGGTTTGTCTATCGGTTTATTTGTTACTGGTTTAAAGAGAGCTTC 31587  
|||||  
4323 QY TTCAGAAATTTGAGTAGATGGAAGAGAGAAAGTAGACATTTCCCTTAAATAATGTACCATGCCT 4382  
|||||  
31588 Db TTCAGAAATTTGAGTAGATGGAAGAGAGAAAGTAGACATTTCCCTTAAATAATGTACCATGCCT 31647  
|||||  
4383 QY TTCAGTTCACAGATCCCTTATCACATGGCTGTCAAGGGTGGCTCAGAAATAGGTAGAGTT 4442  
|||||  
31648 Db TTCAGTTCACAGATCCCTTATCACATGGCTGTCAAGGGTGGCTCAGAAATAGGTAGAGTT 31707  
|||||  
4443 QY AAGAAATTTAAGTAAATTTGGTGAAGCTTAAGCGATGAAGCTTCAATCTTAAAGCTTATATACAT 4502  
|||||  
31708 Db AAGAAATTTAAGTAAATTTGGTGAAGCTTAAGCGATGAAGCTTCAATCTTAAAGCTTATATACAT 31767  
|||||  
4503 QY CAACTGAAATGTAATAATTTGGAACATTTTCCAGGCATCCCTGTTTATTTTATTTTGTCTCT 4562  
|||||  
31768 Db CAACTGAAATGTAATAATTTGGAACATTTTCCAGGCATCCCTGTTTATTTTATTTTGTCTCT 31827  
|||||  
4563 QY CTTTCTCTGCTTGCCTTACTTCAAAAGCTCATATATGGCATGTGAGTGTGACTAGAACTGTGCTGCCAA 4622  
|||||  
31828 Db CTTTCTCTGCTTGCCTTACTTCAAAAGCTCATATATGGCATGTGAGTGTGACTAGAACTGTGCTGCCAA 31887  
|||||  
4623 QY AGAGTTTGTCAATATAGATTCCTTTCTTGTAAACATTTCTACCTTGGGGCTTCAATTTAT 4682  
|||||  
31888 Db AGAGTTTGTCAATATAGATTCCTTTCTTGTAAACATTTCTACCTTGGGGCTTCAATTTAT 31947  
|||||  
4683 QY AATCAAAAGAGTACTGTAACTGTGTCAAAAGCTTACCTGTGACAAATATATATATATGTA 4742  
|||||  
31948 Db AATCAAAAGAGTACTGTAACTGTGTCAAAAGCTTACCTGTGACAAATATATATATGTA 32007  
|||||  
4743 QY TGGTTTACCTGTCAGTAAAGTGGTGGCAATTAATTAATAATATATACAGAAATGAACCGAG 4802  
|||||  
32008 Db TGGTTTACCTGTCAGTAAAGTGGTGGCAATTAATTAATAATATATACAGAAATGAACCGAG 32067  
|||||

```
QY 4803 CAGAACTGTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAATA 4862
|||||
Db 32068 CAGAACTGTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTATACAGTAATA 32127
|||||
QY 4863 ATCAATAAATAGCACTCAATCAAAACATCACTGGGCGACTGTGCACATATATAATC 4922
|||||
Db 32128 ATCAATAAATAGCACTCAATCAAAACATCACTGGGCGACTGTGCACATATATAATC 32187
|||||
QY 4923 AGTGAGATGTGATTGAAGCACAAGGCTTAAGTGAATGCTAGAGAGCTAATTTGATTCAT 4982
|||||
Db 32188 AGTGAGATGTGATTGAAGCACAAGGCTTAAGTGAATGCTAGAGAGCTAATTTGATTCAT 32247
|||||
QY 4983 TTTTATGAAATTTTACTATTATTTAAATGTGCATCCCTGACCATCTTGAACCTTTTACTTGA 5042
|||||
Db 32248 TTTTATGAAATTTTACTATTATTTAAATGTGCATCCCTGACCATCTTGAACCTTTTACTTGA 32307
|||||
QY 5043 AGATTTATTTTATTTTAAATCACTGTTTATTTAGATTTAGTATCTGCGTCTTTGTTTT 5102
|||||
Db 32308 AGATTTATTTTATTTTAAATCACTGTTTATTTAGATTTAGTATCTGCGTCTTTGTTTT 32367
|||||
QY 5103 TCTTTTATCTATGTATGATTTTATTTTATTTTATGAGTGTCTTAAAGCTTCATCAATG 5162
|||||
Db 32368 TCTTTTATCTATGTATGATTTTATTTTATTTTATGAGTGTCTTAAAGCTTCATCAATG 32427
|||||
QY 5163 AGAAGAAATGATTAAATCCATTATTTTACCCCT 5198
|||||
Db 32428 AGAAGAAATGATTAAATCCATTATTTTACCCCT 32463
|||||

RESULT 4
US-10-027-632-100451
; Sequence 100451, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100451
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-100451

Query Match      8.9%; Score 463; DB 13; Length 2057;
Best Local Similarity 80.8%; Pred No. 2e-83;
Matches 612; Conservative 0; Mismatches 115; Indels 30; Gaps 5;

QY 3358 ATATAAGTGTGCCGGAATTTATTTCTTGTGGTTCCTGGTCTCGCTGACTCCAAGA 3417
|||||
Db 783 ATATAAGTTTCTGGAATGGCTCTCTGTGGTCTTGGTCTTCTGCTGACTCAGA 842
|||||
QY 3418 ATGAACCGCAGACCCCTTGTAGTGTCAAGTCTTAAAGATGGTGTGTTCAGAGTT 3477
|||||
Db 843 ATGAACCGCAGACCCCTTGTAGTGTGTAGTCTTAAAGATGGTGTGTGTCCGGAGTT 902
|||||
```

```
QY 3478 TGTTCCTTCAGATGTTCCAGATGTGTCCGAGTGTTCCTCCCTTATGTTGAGTTCGTGTTCTC 3537
|||||
Db 903 TGTTCCTTCAGATGTTCCAGATGTGTCCGAGTGTTCCTCCCTTATGTTGAGTTCGTGTTCTC 962
|||||
QY 3538 GCTGACTTCAACAATGAAGCCGACAGACTTTCAGATGAGTGTGTGACAGTTCCTTAAAGGC 3597
|||||
Db 963 GCTGACTTCAAGAGTGAAGCCACAGACTTCACAGTGA - GTGTCCACAGCTCTTAAAGGT 1020
|||||
QY 3598 AGTGGGTCCAGAGTGTGTTCCTCCGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3657
|||||
Db 1021 GGTACGTCCGAGTGTGTTCCTCCGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
|||||
QY 3658 GAAGTGTCCAGACCCCTCGCGGTAAAGTGTTCAGCTCATAAAGGTAGTGTGCAAAACCCCAACAG 3717
|||||
Db 1081 GAAGTGTCCAGACCCCTCGAGGTGGG - - - - - CAGCTCATAAAGGTAGTGTGCGGACCCCAACAG 1135
|||||
QY 3718 TGACAGTGAAGCAAGATTTATTTATGAAGAGCAAAAGCAAAAGCTTCCCAACATAGAAAC 3777
|||||
Db 1136 TGACAGTGAAGCAAGATTTATTTATGAAGAGTGAAGAGTGAAGAGTTCACAGCTGGAAGG 1195
|||||
QY 3778 GGACCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3833
|||||
Db 1196 GGACCTAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
|||||
QY 3834 ACACCCATCTGCTGATTTGGCCCATTTTACAGAAATGCTGATTGGTCCATTT - - - - - 3886
|||||
Db 1256 CTGCCACATCTGCTGATTGCTC - - - - - TCTTACAGATGCTGATTGCTGCTGCTGCTGCTGCT 1313
|||||
QY 3887 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 3936
|||||
Db 1314 TTTAGCTAGACACAGAGCGCTGATTGGTGTGTTTTTACAGAGTGTGATTGCTGCTGCTGCTGCTGCT 1373
|||||
QY 3937 CAATCCTTTAGCTAGACACAGAGTGTGATTGGTGTGCTTATATCCCTTTAGCTAGACAC 3996
|||||
Db 1374 CAACCTTTAGCTAGACACAGAGTGTGATTGGTGTGTTTTTAAACCTCTAGCTAGACAG 1433
|||||
QY 3997 AAAAGTCTACAAGTCCCAACCCCAAGAGTCCGCTGGCTTCACTCTCGTTAAGGA 4056
|||||
Db 1434 AAAAGTCTCAAGTCCCAACCCCAAGAGTCCGCTGGTGTGTTTTTAAACCTCTAGCTAGACAG 1493
|||||
QY 4057 AATTGAGGTTCAACAAAGTTTCAAGTCTTAAAGTCA 4093
|||||
Db 1494 TTCAATGGCTTGGGTTAGTTTATTACAAACAAATAATA 1530
|||||

RESULT 5
US-10-225-810-26/c
; Sequence 26, Application US/10225810
; Publication No. US20030157512A1
; GENERAL INFORMATION:
; APPLICANT: Birmingham, Jr., John R.
; TITLE OF INVENTION: Tramdorins and Methods of Using Tramdorin
; FILE REFERENCE: McLaugh-07165
; CURRENT APPLICATION NUMBER: US/10/225,810
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 250000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33774)..(33774)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (42953)..(43052)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45557)..(45656)
```



```
QY 3656 ATGAGCTGCAGACCCCTCGCGTAAGTGTACAGCTCATAAAGGTAGTGCAAAACCAAC 3715
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 GTGAAGCCACAGACCTTCGCGGTGAGTGTACAGCTCATAAAGGTAGTGCAAGCCCAAA 817
QY 3716 AGTGAGCAGTAGCAGAGATTATTTGAGAGCAAAAGAACAAAGCTTCCCCACCACATAGAA 3775
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 AGAGAGCAGCAGCAGAGATTATTTGCAAGAGCGAAAGAACAAAGTTTCCACAGCCCTGAA 757
QY 3776 ACGGACCAAGATTGGTTCTCGCTGCTG---TGGTAGCCAGCTTTTATTCCTTATTGG 3831
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 756 GGGGACCAAGACAGCTTCGCGCTGGCTCAGGTGGCCAGCTTTATTCCTTATTGG 697
QY 3832 CCACACCCACATCCTGCTGATTGGGCCCATTTTACAGAATGCTGATTGGTCCATTTATAG 3891
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 696 TCTCGGCC-----CACATCCCGCTGATTGGCCCATTTTACAG 660
QY 3892 CGTCTGATTGGTGGCTTTTACAGAGTGTGATTGGTGCAATTTACAATCCTTTAGCTAG 3951
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 659 AGTGTGGTGTGCTGTGTTTACAAGTGTGATTGGTGCAATTTACAACCTTTAGCTAG 600
QY 3952 ACACAGAGTGTGATTGTGCTTTTATAATCCTTTAGCTAGACACAAAGTCTTACAAGT 4011
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 599 ACACAGAGCGTGTGATTGTGTGTTTACAATCCTTTAGCTAGACAGAAAGTCTCCAAGT 540
QY 4012 CCCCACCAACCCA-GAAGCTCGGCTGCCTTCACCTCTCGTAAGGA 4056
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 CCCCATTGACCCAGGAGTCCAGCTGGCTTCACCTCTCAGCAGGA 494
```

## RESULT 7

```
US-09-728-552-4/c
; Sequence 4, Application US/09728552
; Publication No. US20030096398A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/728,552
; CURRENT FILING DATE: 2000-12-02
; PRIOR APPLICATION NUMBER: 09/078,294
; PRIOR FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; TYPE: DNA
; ORGANISM: Nucleotide sequence of NC-contig
US-09-728-552-4
```

```
Query Match 7.8%; Score 404.6; DB 11; Length 80246;
Best Local Similarity 79.7%; Pred. No. 8.2e-71;
Matches 551; Conservative 0; Mismatches 84; Indels 56; Gaps 4;
```

```
QY 3363 AGTGTGCCCGGAATTTATTTCTCGGTGGGTCTTGCTCGCTCGCTGACTCCAAGATGAA 3422
Db 59807 ACTGTGTGGAATTTGTTCTTCCGTTGGGTCTTGCTCGCTGACTCCAAGATGAA 59748
QY 3423 ACCGAGACCCCTTGAGGTGAGTGTACAGTTCCTTAAAGATGGTGTTCAGAGTTTGTTC 3482
Db 59747 GCCAGACCCCTTGCGGTGAGTGTACAGCTCTGAAGATGGTGTCTGGAATTTGTTC 59688
QY 3483 CTTCAAGATGTTCAAGATGTCCGGAGTTTCTCCCTTATGGTGTGCTCGCTGCA 3542
Db 59687 CTTCAAGATGTTCAAGATGTCTCGGTGGGTCTTGCTCGCTGACTCCAAGATGAA 59628
QY 3543 CTTCAACAAATGAAGCCGAGACCTTTGCAAGTGTGTGACAGTCTTCAAGGCAGTGC 3602
Db 59627 CTTCAAGAGTGAAGCTGACAGCTTCGCAAGTGA--GTGTTACAGCTCTTAAAGGCAGC 59570
QY 3603 GTCCAGAGTTGTTTCTCCCGGTAGTTTCTCGTGGTCTCGCTGATGTCAAGGAATGAAGC 3662
```

```
Db 59569 ATCGGAGTCTTCTTCTCCAGTGGTTCGTGCTCT-GTTGACTTTCAGGAATGAAGC 59511
QY 3663 TGCAGACCCCTCGCGTAAAGTGTACAGCTCATAAAGGTAGTGCAAAACCAACACAGTGAAC 3722
Db 59510 CACAGACCCCTCGCAGTGTAGTGTACAGCTCATAAAGGTAGTGTGGACACACAAGAAATGAC 59451
QY 3723 AGTAGCAAGATTTATTTATGAAGAGCAAAAGCTTCCCCACCATAGAAACGGACCC 3782
Db 59450 AGCAGCAAGATTTATTTGTAAGAGCAAAAGAACAAAGCTTCCACAGTGTGGAAGGGGACC 59391
QY 3783 AGAATGTTGCTGCTGCTG---TGGTAGCCAGCTTTTATTCCTTATTTGGGCCACACC 3838
Db 59390 CGAGTGGGTTCGCCACTGCTGGCTGGGTGGCCAGTTTACTCCCTTATTTGACCCCGCC 59331
QY 3839 CACATCCCTGCTGATTGGCCCATTTTACAGAATGCTGATTGGTCCATTTTATAGCGTGTG 3898
Db 59330 CACATCCCTGCTGATTGGTCCA----- 59310
QY 3899 ATTGGTGGCTTTTACAGAGTGTGATTGGTGCAATTTACAATCCTTTAGCTAGACACAGA 3958
Db 59309 -----TTTACAGAGTGTGATTGGTGCAATTTACAATCCTTTAGCTAGACACAGA 59260
QY 3959 GTGCTGATTGGTGGCTTTTATAATCCTTTAGCTAGACACAAAAGTCTTCAAGTCCCCACC 4018
Db 59259 GTGCTGATTGGTGTGCAATTTACAATCCTTTAGCTAGACACAAAAGTCTTCAAGTCCCCACC 59200
QY 4019 CAACCCCAAGCTCGCTGGCTTCACCTCTC 4049
Db 59199 TGACCCAGAACCCAGCTGGCTTCACCTCTC 59169
```

## RESULT 8

```
US-09-728-552-3/c
; Sequence 3, Application US/09728552
; Publication No. US20030096398A1
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/728,552
; CURRENT FILING DATE: 2000-12-02
; PRIOR APPLICATION NUMBER: 09/078,294
; PRIOR FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Nucleotide sequence of HC-contig
US-09-728-552-3
```

```
Query Match 7.8%; Score 404.6; DB 11; Length 80595;
Best Local Similarity 79.7%; Pred. No. 8.2e-71;
Matches 551; Conservative 0; Mismatches 84; Indels 56; Gaps 4;
```

```
QY 3363 AGTGTGCCCGGAATTTATTTCTCGGTGGGTCTTGCTCGCTCGCTGACTCCAAGATGAA 3422
Db 60040 ACTGTGTCTGGAATTTGTTCTTCCGTTGGGTCTTGCTCGCTGACTCCAAGATGAA 59981
QY 3423 ACCGAGACCCCTTGAGGTGAGTGTACAGTTCCTTAAAGATGGTGTTCAGAGTTTGTTC 3482
Db 59980 GCCACAGACCCCTTGCGGTGAGTGTACAGCTCTGAAGATGGTGTCTGGAATTTGTTC 59921
QY 3483 CTTCAAGATGTTCAAGATGTCCGGAGTTTCTCCCTTATGGTGTGCTCGCTGCA 3542
Db 59920 CTTCAAGATGTTCAAGTGTGTCCAGAGTTCCTTCCCTTGTGGGGTGTGTGGTCTCACTGA 59861
QY 3543 CTTCAACAAATGAAGCCGAGACCTTTGCAAGTGTGTGACAGTCTTAAAGGCAGTGC 3602
Db 59860 CTTCAAGAGTGAAGCTGACAGCTTCGCAAGTGA--GTGTTACAGCTCTTAAAGGCAGCAC 59803
```





[illegible]

```

RESULT 12
US-10-067-514-1
; Sequence 1, Application US/10067514
; Publication No. US20030054531A1
; GENERAL INFORMATION:
; APPLICANT: Gretarsdottir, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reynisdottir, Sigridur Th.
; TITLE OF INVENTION: HUMAN STROKE GENE
; FILE REFERENCE: 2345.2010-003
; CURRENT APPLICATION NUMBER: US/10/067.514
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 09/811/352
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
; TYPE: DNA

```

	Query Match	7.0%	Score 363.6;	DB 14;	Length 1691139;
	Best Local Similarity	80.6%;	Pred. No. 7.4e-62;		
	Matches 464;	Conservative	0;	Mismatches 104;	Indels 8; Gaps 3;
QY	3387	TGTTGGGCTTCTTGCTCTCGCTGACTCCAAAGAATGAAACCGCAGACCCCTTGAGGTGACTGT	3446		
DB	213585	TGTTGGGCTTCTTGCTCTCACTGACTCCAAAGAATCAAGCCCGAGCGCTCATGGTGAAGTGT	213586		





Db 78828 CCACAGTGTTCATTCCTCCAGTGGGTTTGTGCTCTCGCTCGCTCAGGAGTGAAGCTG 78887  
Qy 3665 CAGACCTCGCGGTAAGTGTACAGCTCATAAAGGTAGTGCARACCCCAACAGTGAGCAG 3724  
Db 78888 CAGACCTCGTGGTGAAGTGTACAGCTCATAAAGGCAGTGGGACCCCAAGAGTGATCAA 78947  
Qy 3725 TAGCAAGATTATTATGAAGAGCAAAAGAAAGCTTCCCCACCATAGAAACGACCCAG 3784  
Db 78948 GAGCAAAATTATTGCAAGAGCAAAAGAAAGCTCCACACCATGGAGAGAACCCC 79007  
Qy 3785 AATTGGTTGCTGTGCTG---TGGTAGCCAGCTTTTATTCCTTATTTGGC-CACACCCA 3840  
Db 79008 TGCAGGTTGCTGTGCTGTGGCTGGGCAGCTGCTTTATTCCTTATCTGACTCCACCCA 79067  
Qy 3841 CATCTGCTGATTGGCCCAATTACAGAATTGCTGATTGGTCCCATTTTATAGCGTGCTGAT 3900  
Db 79068 CATCTGCTGATTGGCCCAATTATAGAGAGCTGATTGGTCCCATTTTACAGAGAACTGAT 79127  
Qy 3901 TGGTGCCTTTT-----TACAGAGTGTGATTGGTGCAATTA 3936  
Db 79128 TGGTCCGTTTTACAGAGAGCTGATTGGTCTGTCTCGACAGGGTGTGATTGGTGCAATTA 79187  
Qy 3937 CAATCCCTTTAGCTAGACACAGAGTGTGATTGGTGCCTTTATATCCTTTAGCTAGACAC 3996  
Db 79188 CAATGCCTGAGCTAGACACAGAGTGTGATTGGTGTATTATACAATCCTCTAGCTAGACGT 79247  
Qy 3997 AAAAGTCTACAAGTCCCCAC 4017  
Db 79248 AAAAGTCTCCAAGTCCCCAC 79268

Search completed: September 23, 2003, 06:41:08  
Job time : 1801 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:46:24; Search time 9379 Seconds  
(without alignments)  
13519.187 Million cell updates/sec

Title: US-09-996-956-1

Perfect score: 5217

Sequence: 1 gaactttataatctctc.....taaaaaaaaaaaaaaaa 5217

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estcov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	579.2	11.1	752	10	BF680192
2	497.4	9.5	501	28	AQ345812
3	434.4	8.3	855	12	BM041081
4	392.6	7.5	821	29	BZ600660

C	5	376.8	7.2	600	28	AQ083508
C	6	376.6	7.2	655	29	AG045844
C	7	371.8	7.1	680	29	AG169322
	8	365.2	7.0	649	2	HS078213
	9	362	6.9	607	28	AQ587990
	10	360.8	6.9	845	28	AQ747612
	11	360	6.9	899	13	BQ926428
	12	358	6.9	630	28	AQ390354
	13	354.4	6.8	22715	28	AQ839851
C	14	354.2	6.8	14771	28	AQ839854
	15	352.8	6.8	617	9	AL704833
	16	352.2	6.8	807	10	BF676760
	17	347	6.7	764	28	AQ742559
	18	342.8	6.6	681	29	AG113408
C	19	341.6	6.5	929	14	CD242806
	20	334.2	6.4	674	28	AQ237458
	21	332.2	6.4	680	28	AQ239957
C	22	331.2	6.3	741	29	AG030015
C	23	330.6	6.3	757	10	BG202513
C	24	325.6	6.2	621	28	B72370
C	25	323.8	6.2	846	14	CD171662
C	26	320.2	6.1	907	13	BQ654152
C	27	320.2	6.1	925	13	BQ645350
C	28	314.6	6.0	636	9	AL603245
C	29	314	6.0	984	13	B0839710
C	30	312.4	6.0	555	28	AQ356187
	31	307.2	5.9	854	28	AQ897063
C	32	306.2	5.9	711	28	AQ057655
C	33	304.6	5.8	497	28	AQ481794
C	34	302.6	5.8	666	10	BE393102
C	35	299.8	5.7	652	12	BM978961
C	36	297.2	5.7	406	28	B59220
C	37	297	5.7	460	28	AQ539242
C	38	297	5.7	512	28	B59557
C	39	296.2	5.7	556	28	AQ507503
C	40	295.8	5.7	521	28	AQ075172
C	41	294.6	5.6	460	28	B66805
C	42	293.6	5.6	966	13	BX342527
C	43	292.4	5.6	669	28	B91834
C	44	289	5.5	1914	11	AF289565
C	45	287.6	5.5	547	28	AQ053359

## ALIGNMENTS

BF680192 602154942F1 NIH\_MGC\_83 Homo sapiens CDNA clone IMAGE:4295771 5', linear EST 21-DEC-2000  
 mRNA sequence.  
 BF680192  
 BF680192.1 GI:11954087  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/  
 1 (bases 1 to 752)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue procurement: CLONETECH Laboratories, Inc.  
 CDNA Library prepared by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Plate: LLCML147 row: d column: 12  
 High quality sequence stop: 663.

```
FEATURES
Source
Location/Qualifiers
1..752
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggccgctggcc); Site_2: SfiI (ggccatcatggc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dT(30)BN-3'.
(Where B = A, C, G or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 213 a 168 c 174 g 197 t
ORIGIN
Query Match 11.1%; Score 579.2; DB 10; Length 752;
Best Local Similarity 92.8%; Pred. No. 8.8e-11;
Matches 685; Conservative 0; Mismatches 43; Indels 10; Gaps 7;
QY 173 AATTAGGAGCAACCTTTTTCCTATGCGGTAACTGTGTCTCAACTTCTCTTAATTG 232
Db 2 AATTAGGAGCAACCTTTTTCCTATGCGGTAACTGTGTCTCAACTTCTCTTAATTG 59
QY 233 GGAATAGTTAAGCAGATTCATAGAGCTGAATGATAAAATGTACTAGAGATGCATGG 292
Db 60 GGAATAGTTAAGCAGATTCATAGAGCTGAATGATAAAATGTACTAGAGATGCATGG 119
QY 293 GACTCAAGCTGACCTTATCAAGTCAGATGGAGTCTGCCCTGTCTCAAGGCTGAGGCC 352
Db 120 GACTCAAGCTGACCTTATCAAGTCAGATGGAGTCTGCCCTGTCTCAAGGCTGAGGCC 179
QY 353 AATGGTGTGATCTTGGCTCACTGCAACCTCCAGCTCCAGGTTCAAAAGCTTCTCTGCC 412
Db 180 AATGGTGTGATCTTGGCTCACTGCAACCTCCAGCTCCAGGTTCAAAAGCTTCTCTGCC 239
QY 413 TCAGCTCCCAAGTAAGTGGATTCAGAGGCTTGTGCTGATTTGACACTTCATGATATC 472
Db 240 TCAGCTCCCAAGTAAGTGGATTCAGAGGCTTGTGCTGATTTGACACTTCATGATATC 299
QY 473 AGCCAAAGTGAACATAAAACAGCTCTCTGGAAGAGGACTATGACATCATCAGTTGGGAG 532
Db 300 AGCCAAAGTGAACATAAAACAGCTCTCTGGAAGAGGACTATGACATCATCAGTTGGGAG 359
QY 533 TCTCCAGGACAGCGGACCTTTGGAAAGAGACTAGAAAGTGTGAATCTATTAGTCTTC 592
Db 360 TCTCCAGGACAGCGGACCTTTGGAAAGAGACTAGAAAGTGTGAATCTATTAGTCTTC 418
QY 593 GATATGAATCTCTGCTCTGTAAGAGCTTTCATATTTACAGACACAGGCTACTCC 652
Db 419 GATATGAATCTCTGCTCTGTAAGAGCTTTCATATTTACAGACACAGGCTACTCC 478
QY 653 TAGGCGAGCAAAAAGTGGCAACAGCAGCAGGAGGAGAAA--GAGATCATGAGCATTC 710
Db 479 TAGGCGAGCAAAAAGTGGCAACAGCAGCAGGAGGAGAACACCCGAGATCATGAGCATTC 538
QY 711 AGAGTGCACGTCTTTTCATATATTTCTCAATGCCCTATGTTTGGTTTATTTTGGCCAA 770
Db 539 AGAGTGCACGTCTTTTCATATATTTCTCAAGTGCCTATGTT--GGTTTCATTGGGCCAA 596
QY 771 GCATAACAATCTGCTCAAGAAAATAAATCTGAGAAAACAAGGTCCTTTCGCAATGT 830
Db 597 GCATAGCAATCTGCTCAAGAAAATAAATCTGAGGAAAACAAGGTCCTTTCGCAATGT 656
QY 831 TATGTTCTTTTTCACAGCCCTGAGATTTCTGAGGGGAATTCACATATAATGGATCAGG 890
Db 657 TATGTT--CTTTGACAGGCCCTGAGA--TTCTGAGGGGAATTCACATATAATGGCTCGGT 713
```

```
QY 891 TCATTTCATTACGTTGTG 908
Db 714 CATTCACTTTCGCTGGTG 731

RESULT 2
LOCUS AQ345812
DEFINITION RPCL11-124L14.TV RPCL11 Homo sapiens genomic clone RPCL11-124L14,
genomic survey sequence.
ACCESSION AQ345812
VERSION AQ345812.1 GI:4170708
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
,J.C.
Use of BAC End Sequences from Library RPCL11 for Sequence-Ready
Map Building
Unpublished
Other_GSSs: RPCL11-124L14.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCL11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..501
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7547509"
/db_xref="taxon:9606"
/clone="RPCL11-124L14"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCL11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCL11 Human Male BAC Library"
BASE COUNT 129 a 95 c 121 g 155 t
ORIGIN
Query Match 9.5%; Score 497.4; DB 28; Length 501;
Best Local Similarity 99.6%; Pred. No. 4.4e-08;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3254 TTGAGAGAAATCTGGCAGAGAAACAATAATTTATTTCTTACATAAACCAGCATTTTA 3313
Db 2 TTGAGAGAAATCTGGCAGAGAAACAATAATTTATTTCTTACATAAACCAGCATTTTA 61
QY 3314 CAGGTTTATTTTAACTATGATGTTTATCTGTATCTTTCATATAGTGTGCCGG 3373
Db 62 CAGGTTTATTTTAACTATGATGTTTATCTGTATCTTTCATATAGTGTGCCGG 121
QY 3374 AATTATTTCTTGTGGTCTTGTCTCGCTGACTCCAAGAAATGAACCCGAGACCC 3433
Db 122 AATTATTTCTTGTGGTCTTGTCTCGCTGACTCCAAGAAATGAACCCGAGACCC 181
QY 3434 TTGAGGTGAGTGCACAGTTCTTAAAGATGGTGTTCAGAGTTTGTCTTCAGATGTT 3493
Db 182 TTGAGGTGAGTGCACAGTTCTTAAAGATGGTGTTCAGAGTTTGTCTTCAGATGTT 241
```









```
Matches 524; Conservative 0; Mismatches 98; Indels 54; Gaps 4;
QY 3406 CTGACTCAAGAAATGAACCGCAGACACCTTGAGGTGAGTGC-ACAGTCTCTTAAGATGG 3464
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
680 CTGACTTCAAGATGAAGCGCAGACCCCTGGTGGAGTGTAAACAGTCTCTTAAGATGG 621
QY 3465 TGTGTTCCAGAGTTGTTCCCTTCAGATGTTTCAGATGTCGCGGAGTTCCTCCCTTATGGTG 3524
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
620 TGTCTCCAGAGTTGTTCCCTTCAGATGTTTCAGATGTCGCGGAGTTCCTCCCTTCTGTTG 561
QY 3525 AGTTCGTTGCTCGCTGACTTCAACAATGAAGCGCAGACCTTTGCGAGTGTGAGTGCAC 3584
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
560 GGTGTCGGGCTTCTGACTTCAAGAGTGAAGCTGCAGACCTTCCACAGTGA--GTGTAC 503
QY 3585 AGTTCCTTAAAGGAGTGCCTGCGCAGAGTTGTTGTTCTCCCGGTAGTGTGCTGTCGCG 3644
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
502 AGCTCTTAAAGGTGGTGGTCTGGAGTGTTCATTCCTCTGGTGGTGTGTTGTTCTC 443
QY 3645 TGATGTCAGGAATGAAGTGCAGACCTCGCGTAACTGTTACAGCTCATAAAGGTAGTG 3704
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
442 TGACTTCAGGAGTGAAGCCACAGACCTTCGAGAAAGTGTACAGCTCATAAAGGTAGTG 383
QY 3705 CAACCCAAACAGTGCAGTACGACAGTATTTATGAAGCAAAAGCAAAAGCTTCC 3764
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 CAGACCCAAAGAGTGAAGCAGCATGATTATTTGAAGAGTGAAGAAACAGCTTCC 323
QY 3765 CCACCATAGAAACGAGCAAGATTGTTG--CTGCTGCTGTGTAGCCAGCTTTTATTC 3821
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
322 ACAGCATGAAGGTACCTAGTGGTGGTGGCCCTGCTAGTGGTGGCCAGCTTTTATTC 263
QY 3822 CCTTATTTGGCCACACCATCCTGCTGATTTGGCCCATTTTACAGAAATGCTGATGGTC 3881
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 CCTTATTTGGCCACACCATCCTACTGATTGGTCCA----- 225
QY 3882 CATTTATAGCGTCTGATTGGTGGCTTTTACAGAGTGTGATTTGCTGATTTACAAATC 3941
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
224 -----TTTTTACAGAGTGTGATTTGTTTACAAAC 191
QY 3942 CTTTAGCTAGACAGAGTGTGATTTGGTGCCTTTATAATCCTTTAGCTAGACAAAG 4001
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
190 CTTTAGCTAGACAGAGTGTGATTTGGTGCATTTACATCCTTTAGCTAGACAAAG 131
QY 4002 TTTACAGTCCCAACCAACCAAGCTCCGCTGCTTACCTCTCTGTAAGGAAATG 4061
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 TTCTCAAGTCCCAACCAACCAAGCTCCGCTGCTTACCAAGCTGCTTACCAAGTATGACACTT 71
QY 4062 AGGTTCAACCAAGTTT 4077
DB || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70 AACCTTAACTTACTTT 55
```

```
RESULT 8
HSM078213
ID HSM078213 standard; RNA; EST; 649 BP.
XX AC BX507549;
XX SV BX507549.1
XX
DT 09-MAY-2003 (Rel. 75, Created)
DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKF2p868611625_r1 (from clone DKF2p868611625)
XX EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA
```

```
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (09-MAY-2003) to the EMBL/GenBank/DBJ databases.
BL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY.
XX
CC This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC sequenced by DKFZ (German Cancer Research Center,
CC Heidelberg/Germany) within the cDNA sequencing consortium of
CC the German Genome Project.
CC s1 sequence also available.
CC This clone (DKF2p868611625) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
FH Key Location/Qualifiers
FT source 1. .649
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKF2p868611625"
FT /cclone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
FT DH10B; sites SfiI + SfiIIB"
FT /dev_stage="adult"
FT /tissue_type="CDNA-collection"
XX
SQ Sequence 649 BP; 125 A; 136 C; 186 G; 202 T; 0 other;
Query Match 7.0%; Score 365.2; DB 2; Length 649;
Best Local Similarity 81.8%; Pred. No. 0.00051;
Matches 507; Conservative 0; Mismatches 78; Indels 35; Gaps 6;
QY 3362 AAGTGTCCCGGAATTTATTTCTTCTGGTGGTTCCTCGCTCCGCTCCAGAAATGA 3421
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 AAATGTCTCGGAATTTGTTCTTCTGGTGGTTCCTCGCTCCGCTCCAGAAATGA 91
QY 3422 AACCGCAGACCTTGAGTGTGAGTGTACAGTTCCTTAAAGAGTGTGTTACAGTTCCTT 3481
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
92 AGCGTGGACACTCGCGGTGAGTGTACAGTTCCTTAAAGAGTGTGTTCCGGAATTTGTT 151
QY 3482 CTTTCAGATG-----TTTCAGATGTGTCTCGGAGTTCCTCCCTTATGTGAGTTCGT 3531
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
152 CTTTCAGATGTGTCATTCCTTCAGATGTGTCCAGAGTTCCTTCTCCGCTGGTTCGT 211
QY 3532 GGTCTCGCTGACTTCAACAATGAAGCGCAGACCTTTGCGAGTGTGTGACAGTTCCTT 3591
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GGTCTCGCTGACTTCAAGAGTGAAGCTTACAGCTTCGCGAGT--GGGTGTACAGCTCTT 269
QY 3592 AAAGGCAGTGGTCCAGAGTTGTTGTTCTCCCGGTAGTGTGCTGCTCGCTGATGTC 3651
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 AAAGGTGGTGTGTCGGAGTTCCTTCTCGTGGTGGTCTGCTGCTGCTGCTGCTGCT 329
QY 3652 AGGAATGAAGCTGCAGACCTCGCGGTAAAGTGTGTACAGCTCATAAAGAGTGTGCAACCC 3711
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
330 AGGAACGAAGCTGCAGACCTTCGCGGTGAGTGTGTACAGCTCATAAAGAGTGTGACCC 389
QY 3712 AAACAGTGCAGTAGCAAGATTTATTAAGAGCAAAAGCAAAAGCTTCCGCCCAT 3771
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 AAAGAGTGCAGCAGCAAGATTTATTTGTGAAGAGCAAAAGCAAAAGCTTCCACAGTGT 449
QY 3772 AGAAACGACCAAGATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3827
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 GGAAGGGACCTGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509
QY 3828 TTGGCCACACCCATCCTGCTGATTTGGCCCA--TTTTCAGAAATGCTGATTTGCTCCATTT 3886
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
510 TTGGCCCTGCCCAGCTCCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
QY 3887 -----TATAGCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3929
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
570 ACAACCTTTAGCTAGACAGAGCTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
```



```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM13628 row: g column: 11
High quality sequence stop: 682.

FEATURES
    source
        1..899
            /organism="Homo sapiens"
            /mol_type="mrna"
            /db_xref="taxon:9606"
            /clone="IMAGE:6205474"
            /sex="male"
            /tissue_type="sciatic nerve"
            /dev_stage="adult, 70 yr"
            /lab_host="DH10B"
            /clone_lib="Lupski_sciatic_nerve"
            /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
            NotI; Site_2: SalI; cDNA made by oligo-dT priming.
            Directionally cloned using the following adaptors:
            5'-TCGACCCACCGCTCGG-3' and
            5'-GACATGTTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
            1 kb for average insert length 1.87 kb. This is a primary
            library, non-amplified. Library constructed by Life
            Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
            College of Medicine) and is available through Life
            Technologies."
BASE COUNT      192 a   204 c   241 g   259 t       3 others
ORIGIN
Query Match      6.9%; Score 360; DB 13; Length 899;
Best Local Similarity 78.3%; Pred. No. 0.00058;
Matches 522; Conservative 0; Mismatches 106; Indels 39; Gaps 6;
Qy 3365 TGTGCCCGGAATTTATTTCTTGGTGGGTCTTTGGTCTCGCTGACTCAAGAATGAAC 3424
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
165 TGTGCTGTGGAATTTATTTCTTCCATTTGGTCTTCTCGCTGACTCAAGAATGAAC 224
Qy 3425 CGCAGACCCCTTGAGTCAGTCAGTCACAGTCTTAAGATGGTGTTCAGAGTTTCTCT 3484
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
225 CGCAGACCCCTGCGCAGTGAATTTAAAGCTCTTAAGATGGTGTCTGGAGTTTGTGT 284
Qy 3485 TCAGATGTTTCAGATGTCGCGAGTTTCTCCCTATGTTGAGTTCGCTGCTGACT 3544
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
285 TCAGATGTTTCAGATGTCGCGAGTTTCTCCCTATGTTGAGTTCGCTGCTGACT 344
Qy 3545 TCAACAATGAAGCGCAGACCTTTGTCAGTGTGTGACAGTTCCTTAAAGGCGAGTGGT 3604
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
345 TCGGGAGTGAAGCAGCGGACCTTCGCAGTGA--GTGTTACAGCTCTTAAAGTGGCGGT 402
Qy 3605 CCAGAGTGTGTCTCTCCCGGTAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3664
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
403 CTGGAGTGTGTGTCTCTCCCGAGTGGTGTGT--GTCTGTGCTCACTTGGAGTGAAGCGG 461
Qy 3665 CAGACCCCTCGGCTAAGTGTTCAGAGCTCATAAAGGTAGTGCACAAACCCAGTGCAGCAG 3724
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
462 CAGACCCCTCTGGTGTGAGTGTTCAGAGCTCATAAAGGAGTGGGACCCAGAGTGCAGCAG 521
Qy 3725 TAGCAAGATTTATATGAAGAGCAAAAGACAAAGCTTCCCGACCATAGAACGGACGAG 3784
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
522 CAGCAAGATCTATTGTGAAGAGCGAAAGACAAAGCTTCCACAGCGTGAAGGGGAGCCCC 581
Qy 3785 AATTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3840
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
582 AGCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 641
Qy 3841 CATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3870
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
642 CATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701
Qy 3871 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3929
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

pBACe3.6 vector at EcoRI sites"
BASE COUNT      201 a   184 c   213 g   243 t       4 others
ORIGIN
Query Match      6.9%; Score 360.8; DB 28; Length 845;
Best Local Similarity 79.7%; Pred. No. 0.00058;
Matches 487; Conservative 0; Mismatches 93; Indels 31; Gaps 4;
Qy 3398 TGGTCTCGCTGACTCCAGAAATGAACCGCAGACCCCTTGAGGTGAGTGCACAGTTCTTA 3457
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3 TGATCCGGAAATTCAGAAATGAATCCGAGGACCCCAAGGTGATGTTACAGTCTCTTA 62
Qy 3458 AAGATGCTGTGCTCAGAGTTTGTCTTCAGATGTCAGATGTCGCGAGTTCCTCCCT 3517
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 AAGATGCTGCTCAGAGTTTGTCTTCAGATGTCAGATGTCGCGAGTTCCTCCCT 122
Qy 3518 TATGGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3577
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 TCTGGTGGTTCCTGGTCTTGTGCTGACTTCAGGAGTAAAGCGCAGACCTTTGCATTGA-- 180
Qy 3578 GTGTGACAGTCTTAAAGGAGTCCGTCAGAGTGTGTTGTTCTCTCC--GGTAGTTCGT 3636
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 GTGTGACAGTCTTAAAGAGTGGTGCATCTGCGAGTGTTCATCTCTCCCTGGTGGATTCTG 240
Qy 3637 GGTCTGCTGATGTCAGGAATGAAGCTGCAGACCTCGCGTAAAGTGTTCAGAGTCTATAA 3696
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 GGTCTGCTGACTTCAGAGTGAAGCTGCAGACCTTCGAGTGTATAGTTCATAA 300
Qy 3697 AGGTAGTGCAACCAACAGTGCAGAGTGAAGATTTATTAAGAGCAAAAGAAC 3756
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 AGGTAGTGCAACCAACAGTGCAGAGTGAAGATTTATTAAGAGCAAAAGAAC 360
Qy 3757 AAGTCTCCCAACCATGAACAGCAGCAATGTTGTTCTGCTGCTGCTGCTGCTGCTGCTG 3812
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 AAGTCTCCCAACCATGAACAGCAGCAATGTTGTTCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 3813 CTTTATTCCTTATTTGGCCACACCCACATCTCTG-----C 3848
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 CTTTATTCCTTATTTGGCCCGCCACATCTCTGTTCTTGGTCCATTTTACAGAGCAC 480
Qy 3849 TGATGGCCCATTTACAGAAATGCTGATTTGGTCCATTTTATAGGCTGCTGATGGTGGCT 3908
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 TGATGGTCCATTTACAGAGTGTGATTTGGTCCATTTTACAGAGTCTGATCGTGGGA 540
Qy 3909 TTTTACAGAGTGTGATTTGGTGCATTTACAACTCTTTAGCTAGACAGAGTGTGCTGATG 3968
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 TTTTACAGAGTGTGATTTGGTGTGTTTACAAACCGTTAGCTAGACAGAGCAGTGTG 600
Qy 3969 GTGCTTTTATA 3979
Db      ||| ||| |||
601 GTGCGTNTTTA 611

RESULT 11
BQ926428
LOCUS
DEFINITION
AGENCOURT_8754346 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6205474 5', mRNA sequence.
ACCESSION
BQ926428
VERSION
BQ926428.1 GI:22341459
KEYWORDS
EST.
SOURCE
Homo sapiens
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 899)
NTH-MGC http://mnc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
```

```

Db      702 GCTGATTGGTCCATTTTACAGAGCGCTGAGTGGTGGTTTTTACACAGTGGCGTGATTGCT 761
QY      3930 GCATTACAAATCCTTTAGCTACACAGAGTGCT-GATTGGTGGCTTTATATACCTTTAG 3988
Db      762 GCATTANCATCCTTTAGCTAGACCAGAGTGCTNGAATGGTGCAATTTTACATACATTTAG 821
QY      3989 CTAGACA 3995
Db      822 ACACAGA 828

RESULT 12
LOCUS   AQ390354
DEFINITION CITBI-E1-2545M11.TR CITBI-E1 Homo sapiens genomic clone 2545M11.
ACCESSION AQ390354
VERSION   AQ390354.1 GI:4361377
SOURCE   GSS.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 630)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
JOURNAL Map Building
COMMENT Unpublished
Other_GSSs: CITBI-E1-2545M11.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..630
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /clone="2545M11"
     /sex="male"
     /cell_type="sperm"
     /clone_lib="CITBI-E1"
     /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT          144 a 134 c 175 g 177 t
ORIGIN

Query Match          6.9%; Score 358; DB 28; Length 630;
Best Local Similarity 83.0%; Pred. No. 0.00088;
Matches 448; Conservative 0; Mismatches 80; Indels 12; Gaps 3;

QY      3356 TCATATAAGTGGCCCGGAATTTATTTCTCTGGTGGGTTCTTGCTCGCTGACTCCAA 3415
Db      98 TCACCAAAATGTGTCGGGAATTTATCTCTCCAGTGAGTTTACAGCTCTTAAAGATGGTGTGCGGAG 157
QY      3416 GAATGAAACCGCAGACCCCTTGAGTGAGTGTCACAGTTCTTTAAAGATGGTGTGTTTCAGAG 3475
Db      158 GAATGAAGCCAGGACCCCTTGACGTGAGTGTGTACAGCTCTTAAAGATGGTGTGTCGGAG 217
QY      3476 TTGTGTCCTTCAGATGTTTCAGATGTCGGAGTTCTCCCTTATGGTGAGTTCTGTGTC 3535
Db      218 TTGTGTCCTTCAGAGATTCAGATGTCATCCAGAGTTTCTTCCCTCTGTTGGTGGTGTGTC 277
QY      3536 TCGCTGACTTCAACAATGAAGCCGACACCTTTGTCAGTGAGTGTGTGACAGTCTCTTAAAG 3595

Db      278 TTGTGAGCTTCAGGAGTGA-----AGACTTTGCAGTGA--GTGTACAGCTCTTAAAG 329
QY      3596 CGAGTGGCTGCAGAGATTGTTGTTCTCCCGTAGTGGTTCGTCGTCGTCGTCGTCGTCGTCGTCG 3655
Db      330 GTGGTGGCTGCGGAGTTGTTTGTCTCCCGTGGGTTCGTCGTCGTCGTCGTCGTCGTCGTCG 389
QY      3656 ATGAAGCTGCAGACCCCTCGCGTAAAGTGTTCACAGCTCATAAAGTAGTCAAAACCCCAAC 3715
Db      390 ATGAAGCCACAGACTCTTTCGAGTGAGTGTACAGCTCATAAAGTAGTTCGCGACCCCAAG 449
QY      3716 AGTGAGCAGTAGCAAGATTATATGAAGAGCAAAAGACAAAGCTTCCCCACCATAGAA 3775
Db      450 AATGAGCAGCACCAAGATTATATGAAGAGCAAAAGACAAAGCTTTCAAAGCGTGGAA 509
QY      3776 ACGGACCAAGATTGCTGCTGCT---GCTGTGGTAGCCAGCTTTTATCCCTTATTTGG 3831
Db      510 GGGGACCCGAGCGGGTGGCCACTGCTGGAGTGGCCAGCTTTTATCCCTTATTTCT 569
QY      3832 CCACACCCACATCCTGCTGATTGGCCCAATTTTACAGAAATGCTGATTGGTCCATTTATAG 3891
Db      570 CCCCGTCCACATCCTGCTGATTGGTCCATTTTACAGAGTGTGATTGGTCCATTTTCATAG 629

RESULT 13
LOCUS   AQ839851
DEFINITION 69115-C79 CITB Homo sapiens genomic clone 69115, genomic survey
sequence.
ACCESSION AQ839851
VERSION   AQ839851.1 GI:6652483
KEYWORDS GSS.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 22715)
AUTHORS Carpten,J.D., Makalowski,I., Robbins,C.M., Scott,N., Sood,R.,
Connors,T.D., Bonner,T.I., Smith,J.R., Faruque,M.U., Stephan,D.A.,
Pinkett,H., Morgenbesser,S.D., Su,K., Graham,C., Gregory,S.G.,
Williams,H., McDonald,L., Baxeavanis,A.D., Klingler,K.W. and Landes
,G.M.
TITLE A 6-Mb high-resolution physical and transcription map encompassing
the hereditary prostate cancer 1 (HPC1) region
JOURNAL Genomics 64 (1), 1-14 (2000)
MEDLINE 20175426
PUBMED 10708513
COMMENT Contact: Carpten JD
Cancer Genetics Branch
National Human Genome Research Institute/National Institutes of
Health
Bldg. 36, Room 3D04, 36 Convent Drive, Bethesda, MD
Tel: 301 435 5626
Fax: 301 435 5465
Email: jdc@nhgri.nih.gov
Class: Shotgun.

FEATURES             Location/Qualifiers
     source            1..22715
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /clone="69115"
     /clone_lib="CITB"
BASE COUNT          7023 a 4136 c 4231 g 7322 t 3 others
ORIGIN

Query Match          6.8%; Score 354.4; DB 28; Length 22715;
Best Local Similarity 82.2%; Pred. No. 7.9e-05;
Matches 457; Conservative 0; Mismatches 91; Indels 8; Gaps 4;

QY      3504 CGGAGTTTCTCCCTTATGGTGGTTCGCTGACTTCAACAATGAAGCCGCGAGA 3563
Db      20272 CTGGATTGGTTCCTCCAGTGGTTCCTGCTCACTGACTTCAAGAGTGAAGCCGCTGA 20331

```

```
QY 3564 CCTTTGACGTGAGTGTGTGACAGTCTTAAAGCAGTGCCTCCAGAGTGTGTTGTCCTC 3623
Db 20332 CCCTCGCGGTGA--GTGTACAGTCTTAAAGATGGCGCGTCTGGAGTTTTTTGTTCTC 20389
QY 3624 CCGGTAGTGTGCTGCTCGC--TGATGTCAGGAATGAAGCTGCA--GACCCCTCGCGGTAAAG 3681
Db 20390 CCAGTGGGTTGTGCTCGCTGACCTCAGAGTGAAGCGCAGGACCTTCGCACATCA 20449
QY 3682 TGTACAGCTCATAAAGTAGTGCACAAACCCAAACAGTACAGTACGAAGATTATTATG 3741
Db 20450 TGTACAGCTCATAAAGTAGTGCACAGACCCAAACAGTACAGCAGCAAGATTACGGTG 20509
QY 3742 AAGAGCAAAAGCAAAAGCTTCCCAACCATAGAACGACACAGATTGGTGTGCTGCT 3801
Db 20510 AAGAGCAAAAGCAAAAGCTTCCACAGCAGCGAAGGGACCGCAGTACTGCTGCCACTGCT 20569
QY 3802 G----TGGTAGCCAGCTTTTATTCCTTATTGCGCCACACCCACATCCTGCTGATGGCC 3857
Db 20570 GGCTCGGTAGCCAGCTTTTATTCCTTATTGCGCCCTGCCCATGTACTGCTGATGGTC 20629
QY 3858 CATTTTACAGAACTGCTGATGGTCCATTTATACGGTGTGATGGTGGTGGTTCACAGA 3917
Db 20630 CATTTTACAGAGCCTGATGGTCCATTTTACAGAGTCTGATGGTGGTGGTTCACAGA 20689
QY 3918 GTGCTGATGGTGCATTTTAAATCCTTTAGTAGACACAGAGTCTGCTGATGGTGGTCCCTTA 3977
Db 20690 GTGCTGATGGTGCATTTTAAATCCTTTAGTAGACACAGAGTCTGCTGATGGTGGTGGT 20749
QY 3978 TAATCCTTTAGTAGACACAAAGTCTTACAAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGT 4037
Db 20750 CAATCCTTTAGTAGACACAAAGTCTTACAAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGT 20809
QY 4038 GCTTCACCTCTCGTAA 4053
Db 20810 GCTTCACCTCTCTCTTA 20825

RESULT 14
LOCUS A0839854 14771 bp DNA linear GSS 31-MAR-2000
DEFINITION 260L13-C54 CITB Homo sapiens genomic clone 260L13, genomic survey
sequence.
ACCESSION A0839854
VERSION A0839854.1 GI:6652486
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 14771)
Carpen, J.D., Makalowska, I., Robbins, C.M., Scott, N., Sood, R.,
Connors, T.D., Bonner, T.I., Smith, J.R., Faruque, M.U., Stephan, D.A.,
Pinkett, H., Morgenbesser, S.D., Su, K., Graham, C., Gregory, S.G.,
Williams, H., McDonald, L., Baxevanis, A.D., Klingler, K.W. and Landes
, G.M.
TITLE A 6-Mb high-resolution physical and transcription map encompassing
the hereditary prostate cancer 1 (HPC1) region
JOURNAL Genomics 64 (1), 1-14 (2000)
MEDLINE 20175426
PUBMED 10708513
COMMENT Contact: Carpen JD
Cancer Genetics Branch
National Human Genome Research Institute/National Institutes of
Health
Bldg. 36, Room 3D04, 36 Convent Drive, Bethesda, MD
Tel: 301 435 5626
Fax: 301 435 5465
Email: jdc@nhgri.nih.gov
Class: shotgun
Location/Qualifiers
1. .14771
/organism="Homo sapiens"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="CITB"
BASE COUNT 4135 a 2762 c 2928 g 4937 t 9 others
ORIGIN
Query Match 6.8%; Score 354.2; DB 28; Length 14771;
Best Local Similarity 77.08; Pred. No. 0.00011;
Matches 532; Conservative 0; Mismatches 103; Indels 56; Gaps 6;
QY 3358 ATATAAGTGTCCCGGGAATTTATTTCTTCTGGTGGTTCCTGGTCTGCTGACCTCCAGA 3417
Db 2239 ATAACATTTGTCCAGAGTTTATCCCTTCCAGTGGTTCCTGCTCGCTGACITCAAGA 2180
QY 3418 ATGAACCGGAGACCCCTTGAGTGTGACAGTCTTAAAGATGGTGTGTCAGAGTT 3477
Db 2179 ACGAAGCGGAGAACTTCAAGTGTGACAGTCTTAAAGATGGTGTGTCGGGAAT 2120
QY 3478 TGTTCCTTCAGATGTCAGATGTCGCGAGTTCCTCCCTTATGCTGAGTTCGTTGCTC 3537
Db 2119 TGTTCCTTCAGAGTGTCTA-----GGAGTTTCTTCCCTCCGCTGAGTTAGTGGTCTC 2067
QY 3538 GCTGACTTCAACATGAAGCGCAGACCTTTGACGTGAGTGTGACAGTTCCTTAAAGGC 3597
Db 2066 GCTGACTTCAGGAGTGAAGCTGTGACCTCCAGTGTGAGT--GTTACAGCTCTTAAAGGT 2009
QY 3598 AGTCGCTCCAGAGTGTGTTGTTCTCTCCGGTAGTTCGTGCTGCTGCTGATGTCAGGAT 3657
Db 2008 GGCGCTCTGAGAGTGTGTTGGCCCTCTGCTGGGTTCATGCTGCTGCTGCTGAGTTCAGGAAT 1949
QY 3658 GAAGCTCGACAGCCCTCGGCTAAGTGTACAGCTCATAAAG---GTAGTGCAAAACCCAAA 3714
Db 1948 GAAGCTCGACAGCCCTCGGCTGAGTGTGACAGTCTATAAGGTAGTAGTCCGACCCAAA 1889
QY 3715 CAGTGACAGTACAGAGATTTATTAAGAGCAAAAGAAAGAGTTCCTCCACCATAGAG 3774
Db 1888 GAGTGACAGCAGCAAGATTTATTTGAGGAGCAAAACAAAGAGTTCACACATAGA 1829
QY 3775 AACGGACACAGATTTGTTGCTGCTG---CTGTGGTAGCCAGCTTTTATCCCTTATTTG 3830
Db 1828 AGTGGACAGAGAGAGCTGCCACTGGGCGCTTCGGTGGCCAGCTTTTATCCCTTATTTG 1769
QY 3831 GCCACACCCACATCCTGCTGATGGCCCATTTTACAGATGCTGATGGTGCATTTAT- 3889
Db 1768 GCCCTGCCACATCCTACTGATGGTCCATTTTACAGAGTCTGATGGTGCATTTTTC 1709
QY 3890 AGCGTGTGATGGTGGCTG---T-----T 3910
Db 1708 ACAGTGTGACTGTGCTGATTTTACAATCCTTTAGCTAGACACAGAGCAGTGAATGGTGCAT 1649
QY 3911 TTACAGAGTCTGATGGTGCATTTACAATCCTTTAGCTAGACACAGAGTCTGATGGT 3970
Db 1648 TTTTAGAGTCTGACTGGTGCATTTTACAATCCTTTAGCTAGACACAGAGGCTGATGGT 1589
QY 3971 GCCTTTATAATCCTTTAGCTAGACACAAAG 4001
Db 1588 AAATTTACAATCCTTTAGCTAGACAAAGAG 1558

RESULT 15
LOCUS AL704833
DEFINITION DREF2686B2033_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
AL704833
ACCESSION AL704833
VERSION AL704833.1 GI:19688188
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 617)
```





**THIS PAGE BLANK (2570)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2003, 17:01:44 ; Search time 18295 Seconds  
(without alignments)  
11665.784 Million cell updates/sec

Title: US-09-996-956-1  
Perfect score: 5217  
Sequence: 1 gaaacttaaaatccctc.....taaaaaaaaaaaaaaaa 5217

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	5217	100.0	5217	6	AX467721	Sequence
2	4957	95.0	5092	6	AX467723	Sequence
3	4763.6	91.3	168595	9	AC013401	Homo sapi
4	4758.4	91.2	32463	6	AX467725	Sequence
5	1458.4	28.0	1603	9	AF223389	Homo sapi
6	514.8	9.9	131458	9	AC002386	Homo sapi
7	510.4	9.8	193041	9	AC092017	Homo sapi
8	503	9.6	162126	9	AL354711	Human DNA
9	503	9.6	297235	2	AL499603	Homo sapi
10	494.6	9.5	99632	9	AL157779	Human DNA
11	484.4	9.3	123637	9	HS3067	Human DNA
12	484.2	9.3	99527	9	AL137015	Human DNA
13	484.2	9.3	135090	9	HS287614	Human DNA
14	483.8	9.3	159797	9	AC104697	Homo sapi
15	482.4	9.2	179598	9	CNS01DRO	Human chr
16	482	9.2	46975	9	AL359376	Human DNA
17	482	9.2	68697	9	AL451009	Human DNA
18	480.4	9.2	145993	9	AC098972	Homo sapi
19	480.4	9.2	158508	9	AC107622	Homo sapi
20	479.4	9.2	159851	9	AC036114	Homo sapi
21	479.4	9.2	164009	9	AC079064	Homo sapi
22	479.4	9.2	195344	2	AC025248	Homo sapi
23	479.2	9.2	76028	9	AL451129	Human DNA
24	477.8	9.2	117583	9	AC034238	Homo sapi
25	477.8	9.2	133322	2	AC091845	Homo sapi
26	476.4	9.1	203076	9	AL591074	Human DNA
27	476.2	9.1	183494	9	AC091977	Homo sapi
28	470.8	9.0	103728	9	AL445219	Human DNA
29	470.4	9.0	3265	9	HSAC001178	Homo sapi
30	470.4	9.0	62466	9	AC005613	Homo sapi
31	468.4	9.0	188818	2	AC098674	Homo sapi
32	467.4	9.0	144626	9	AL159159	Human DNA
33	467	9.0	114016	9	AL365396	Human DNA
34	466.8	8.9	52813	9	AC010991	Homo sapi
35	465.8	8.9	32213	2	AC012120	Homo sapi
36	465.8	8.9	126253	9	AP002023	Homo sapi
37	465.8	8.9	127143	2	AL138736	Homo sapi
38	465.8	8.9	173933	9	AP002022	Homo sapi
39	465.2	8.9	222883	2	AC092847	Homo sapi
40	465	8.9	142234	9	AC110620	Homo sapi
41	465	8.9	171196	2	AP001369	Homo sapi
42	463.2	8.9	157604	9	AC021093	Homo sapi
43	463.2	8.9	157633	9	AC123776	Homo sapi
44	463	8.9	108182	9	AC009308	Homo sapi
45	462	8.9	133614	2	AC025433	Homo sapi

ALIGNMENTS

RESULT 1  
AX467721  
LOCUS AX467721 5217 bp DNA linear PAT 16-JUL-2002  
DEFINITION Sequence 1 from Patent WO0244332.  
ACCESSION AX467721  
VERSION AX467721.1 GI:21900893  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Sun, Z., Li, X. and Jay, G.  
TITLE Prostate polynucleotides and uses  
JOURNAL Patent; WO-0244332-A-1-06-JUN-2002;

FEATURES		Origene Technologies Inc. (US)	
Source	1..5217	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	1588 a	972 c	988 g 1669 t
ORIGIN			
Query Match	100.0%;	Score 5217;	DB 6; Length 5217;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 5217;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	GAACACTTTAAATATCCCTCAGTCTCTGTTAAATTCATGTAGTGCCTCCCAAGGCACCTCT	60
Db	1	GAACACTTTAAATATCCCTCAGTCTCTGTTAAATTCATGTAGTGCCTCCCAAGGCACCTCT	60
QY	61	GGCACCCAGTTTGGAACTGCAGTTTAAAAGTCATAAATGAAATGAAATGATAGCAAA	120
Db	61	GGCACCCAGTTTGGAACTGCAGTTTAAAAGTCATAAATGAAATGAAATGATAGCAAA	120
QY	121	GGTGAGGTTTAAAGAGCTATTATAGTCCCTGGACAGCATCTTTTCAATTAGGC	180
Db	121	GGTGAGGTTTAAAGAGCTATTATAGTCCCTGGACAGCATCTTTTCAATTAGGC	180
QY	181	AGCAACCTTTTGGCCCTATGCCGTAACCTGTGTGCAAACTTCTCTAAATGGGAAATAG	240
Db	181	AGCAACCTTTTGGCCCTATGCCGTAACCTGTGTGCAAACTTCTCTAAATGGGAAATAG	240
QY	241	TTAAGCAGATTATAGAGCTGAATGATGATAAATTTGCTACGATGACGCTGGGACTCAAC	300
Db	241	TTAAGCAGATTATAGAGCTGAATGATGATAAATTTGCTACGATGACGCTGGGACTCAAC	300
QY	301	GTGACCTTATCAAGTGAAGTGGAGTCTTGCCCTGCCCTCCCAAGGCTGGAGCCCAATGGTGT	360
Db	301	GTGACCTTATCAAGTGAAGTGGAGTCTTGCCCTGCCCTCCCAAGGCTGGAGCCCAATGGTGT	360
QY	361	GATCTTGGCTCAGTGCACCTCCACCTCCAGGTTCAAAAGTTTCTGCTCCAGGCTC	420
Db	361	GATCTTGGCTCAGTGCACCTCCACCTCCAGGTTCAAAAGTTTCTGCTCCAGGCTC	420
QY	421	CCAAAGTAACCTGGGATACAGAGGCTTGGTGCAATTTGACACTTCAATGATCAGCCAAAG	480
Db	421	CCAAAGTAACCTGGGATACAGAGGCTTGGTGCAATTTGACACTTCAATGATCAGCCAAAG	480
QY	481	TGAACTTAAACAGCTCCTGGAAGAGGACTATGACATCATCAGGTTGGAGTCTCCAGG	540
Db	481	TGAACTTAAACAGCTCCTGGAAGAGGACTATGACATCATCAGGTTGGAGTCTCCAGG	540
QY	541	GACAGCGACCCCTTTGGAAAGGACTAGAAAGTGTAATCTATTAGTCTTCGATATGAA	600
Db	541	GACAGCGACCCCTTTGGAAAGGACTAGAAAGTGTAATCTATTAGTCTTCGATATGAA	600
QY	601	ATTCTCTGCTCTGTAAGACATTTCAATTTACAGACACAGCCCTACTAGGCGAG	660
Db	601	ATTCTCTGCTCTGTAAGACATTTCAATTTACAGACACAGCCCTACTCTAGGCGAG	660
QY	661	CAAAAGTGCACACAGCAAGCAGAGGAAAGAGATCATGAGGCATTTTCAGAGTGCAC	720
Db	661	CAAAAGTGCACACAGCAAGCAGAGGAAAGAGATCATGAGGCATTTTCAGAGTGCAC	720
QY	721	GTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTATTTTGGCCCAAGCATCAAA	780
Db	721	GTCTTTTCATATATTTCTCAATGCCGTATGTTGGTTTATTTTGGCCCAAGCATCAAA	780
QY	781	CTGCTCAAGAAAAAATCTGGAAAAACAAGGTGCCCTTGGCCAAATGTTATGTTCTT	840
Db	781	CTGCTCAAGAAAAAATCTGGAAAAACAAGGTGCCCTTGGCCAAATGTTATGTTCTT	840
QY	841	TTTGACAGCCCTCAGATTCTCAGGGGAATTCACATAAATGGATCAGTCAATTT	900
Db	841	TTTGACAGCCCTCAGATTCTCAGGGGAATTCACATAAATGGATCAGTCAATTT	900

Qy	901	ACGTTGTGTCAAAATATGATTTAAAGATACAACTTTGCAGAGAGCAGCTTTCTCTAAGG	960
Db	901	ACGTTGTGTCAAAATATGATTTAAAGATACAACTTTGCAGAGAGCAGCTTTCTCTAAGG	960
Qy	961	GTAGGCAGTGGAGGACTAAGGTAAGCATTTCTTAAGATCAGTTAATCAAGAAAGGTG	1020
Db	961	GTAGGCAGTGGAGGACTAAGGTAAGCATTTCTTAAGATCAGTTAATCAAGAAAGGTG	1020
Qy	1021	CTCTTTGCATTTGAAATGCCCTTTGTTGCAAAATATGTTGTTAATTAATTTACACT	1080
Db	1021	CTCTTTGCATTTGAAATGCCCTTTGTTGCAAAATATGTTGTTAATTAATTTACACT	1080
Qy	1081	TAATGGAACAACTTTAACTTACAGATGAACAAACCCACAAAGCAAAATCAAAAGC	1140
Db	1081	TAATGGAACAACTTTAACTTACAGATGAACAAACCCACAAAGCAAAATCAAAAGC	1140
Qy	1141	CCTACCTATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTGCTTTGG	1200
Db	1141	CCTACCTATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTGCTTTGG	1200
Qy	1201	GCATAAATGATTAAGGATTTTCCAGGTATTTTAAATGAGGCCCATCTACAAT	1260
Db	1201	GCATAAATGATTAAGGATTTTCCAGGTATTTTAAATGAGGCCCATCTACAAT	1260
Qy	1261	CTTAGCAATACCTTTGGGATAATTTAAATTCAGCTGGACATTTGCTAATGTTTTATA	1320
Db	1261	CTTAGCAATACCTTTGGGATAATTTAAATTCAGCTGGACATTTGCTAATGTTTTATA	1320
Qy	1321	TACATCTTTGCTAGAAATTTCAATTTTAAATGATGTAATTTAGTTAGTGTGCTGA	1380
Db	1321	TACATCTTTGCTAGAAATTTCAATTTTAAATGATGTAATTTAGTTAGTGTGCTGA	1380
Qy	1381	TCAATTCAAAACAACTTACTTTCCATAATTTTAGACTATGAAGGTCATAAATCAACAAT	1440
Db	1381	TCAATTCAAAACAACTTACTTTCCATAATTTTAGACTATGAAGGTCATAAATCAACAAT	1440
Qy	1441	ATATCTACATCAATATAGATTTTTCATTATAATGCTCTTCATCTTAACAGAAAT	1500
Db	1441	ATATCTACATCAATATAGATTTTTCATTATAATGCTCTTCATCTTAACAGAAAT	1500
Qy	1501	GTCTTTGTGATTTTGTAGAAAACCTGAGAGTTTAAATTCATAATTTACTTGATCAAAA	1560
Db	1501	GTCTTTGTGATTTTGTAGAAAACCTGAGAGTTTAAATTCATAATTTACTTGATCAAAA	1560
Qy	1561	TTGTGGAAACAATCCAGCATTAATTTGATGTTGTTTATGTACATAAGGAGTCTTA	1620
Db	1561	TTGTGGAAACAATCCAGCATTAATTTGATGTTGTTTATGTACATAAGGAGTCTTA	1620
Qy	1621	AGCTTGTGCTTGAAGTCTTTTGTACTTAGTCCCTGTTTAAAATTTACTACTTATATC	1680
Db	1621	AGCTTGTGCTTGAAGTCTTTTGTACTTAGTCCCTGTTTAAAATTTACTACTTATATC	1680
Qy	1681	TAAAGCATTTTATGTTTTCATTTCAATTTACATGATGCTAATTTATGCAATTAACAAA	1740
Db	1681	TAAAGCATTTTATGTTTTCATTTCAATTTACATGATGCTAATTTATGCAATTAACAAA	1740
Qy	1741	TATTAAGATTTTCAAAATAGAATATGTGTAATTTGTTTCACATACATAGAAATGAAAGTCA	1800
Db	1741	TATTAAGATTTTCAAAATAGAATATGTGTAATTTGTTTCACATACATAGAAATGAAAGTCA	1800
Qy	1801	TTTGTGTAAGCAAGATGCTGGGTGAAAGAGTGTCTTTGATTGAAAGATCACTAGTAGT	1860
Db	1801	TTTGTGTAAGCAAGATGCTGGGTGAAAGAGTGTCTTTGATTGAAAGATCACTAGTAGT	1860
Qy	1861	AGAGGCAAGACTTCTAGTCCCTAATCTACCTTAATAGCCATGTCAGCTGAAGTC	1920
Db	1861	AGAGGCAAGACTTCTAGTCCCTAATCTACCTTAATAGCCATGTCAGCTGAAGTC	1920
Qy	1921	AGTGAACCCCTCTCATCTCTCTCATCTTTTTCATCTCTTAAATGAGGGTATATTTAA	1980
Db	1921	AGTGAACCCCTCTCATCTCTCTCATCTTTTTCATCTCTTAAATGAGGGTATATTTAA	1980
Qy	1981	GCTCTTCATTTTCTTTTCTTTTGTGATAGAGTCTTGTCTCTGTACCCAGGTTGGAGTG	2040









Db	2756	TTCTTCTTACTCAGAGATTTTTCGAGAGGTAATGTAATGATCCAGAGGTGAGAATCACTG	2815
Qy	2941	CCTTTTTAATGCGATTTAAATTTGATGAACAAAGTTTCCAAACAAATAACAGTAATAAAA	3000
Db	2816	CCTTTTTAATGCGATTTAAATTTGATGAACAAAGTTTCCAAACAAATAACAGTAATAAAA	2875
Qy	3001	GAACATGTAATTAGCACTTAATAAGCCAGGGCTGGAGAGCTGTGTACATGCTTTCAA	3060
Db	2876	GAACATGTAATTAGCACTTAATAAGCCAGGGCTGGAGAGCTGTGTACATGCTTTCAA	2935
Qy	3061	TCCATGAACCTGGTAAACTGGTACTAGTATCTCTATTGGACATGTGAGAAACCAATGGA	3120
Db	2936	TCCATGAACCTGGTAAACTGGTACTAGTATCTCTATTGGACATGTGAGAAACCAATGGA	2995
Qy	3121	GTTCATTAACAGTAGAGTTAAAAATTAATCTTTCATATATATATATGCTCAATCTCACAG	3180
Db	2996	GTTCATTAACAGTAGAGTTAAAAATTAATCTTTCATATATATATATGCTCAATCTCACAG	3055
Qy	3181	ACATCTCTGCTACCAAAAGCTATCATATCTAGATATGCGCATAGAGTACCTTGGGCG	3240
Db	3056	ACATCTCTGCTACCAAAAGCTATCATATCTAGATATGCGCGATAGAGTACCTTGGGCG	3115
Qy	3241	ACACTAGAAATCTTTTGAGAGAAATCTGCGAGAGAAAACAAATATTTATCTCAATAAAA	3300
Db	3116	ACACTAGAAATCTTTTGAGAGAAATCTGCGAGAGAAAACAAATATTTATCTCAATAAAA	3175
Qy	3301	ACCCAGCATTTTACAGGTTTATTTTAACTATGAAGTATGTTATCTGTATCTTTTATA	3360
Db	3176	ACCCAGCATTTTACAGGTTTATTTTAACTATGAAGTATGTTATCTGTATCTTTTATA	3235
Qy	3361	TAAGTGTGCCCGGAATTTATTTCTCTGGTGGGTTCTTGGTCTCGCTGACTTCCAAGAAATG	3420
Db	3236	TAAGTGTGCCCGGAATTTATTTCTCTGGTGGGTTCTTGGTCTCGCTGACTTCCAAGAAATG	3295
Qy	3421	AAACCCGAGACCCCTTGAGGTGAGTGCACAGTCTTAAAGATGGTGTTCAGAGTTTGT	3480
Db	3296	AAACCCGAGACCCCTTGAGGTGAGTGCACAGTCTTAAAGATGGTGTTCAGAGTTTGT	3355
Qy	3481	TCCTTCAGATGTTTCAGATGTCCGGAGTTTCTCCCTTATGGTGAGTTCGTGGTCTCGCT	3540
Db	3356	TCCTTCAGATGTTTCAGATGTCCGGAGTTTCTCCCTTATGGTGAGTTCGTGGTCTCGCT	3415
Qy	3541	GACTTCAACAAATGAAGCCGAGACCTTTTGCACTGAGTGTGTGACAGTTCTTAAAGGCAGT	3600
Db	3416	GACTTCAACAAATGAAGCCGAGACCTTTTGCACTGAGTGTGTGACAGTTCTTAAAGGCAGT	3475
Qy	3601	CGCTCCAGAGTTGTTTGTCTCCCGGTAGTTCGTGGTCTCGCTGATGTACAGAAATGAA	3660
Db	3476	CGCTCCAGAGTTGTTTGTCTCCCGGTAGTTCGTGGTCTCGCTGATGTACAGAAATGAA	3535
Qy	3661	GCTGCAGACCTCGCGGTAAGTGTACAGCTCATAAAGGTAGTGCAAACCCAAACAGTGA	3720
Db	3536	GCTGCAGACCTCGCGGTAAGTGTACAGCTCATAAAGGTAGTGCAAACCCAAACAGTGA	3595
Qy	3721	GCAGTAGCAAGATTTATTAAGAGCAAAAAGCAAAAGCTTCCCCACCATAGAAACGGA	3780
Db	3596	GCAGTAGCAAGATTTATTAAGAGCAAAAAGCAAAAGCTTCCCCACCATAGAAACGGA	3655
Qy	3781	CCAGAAATGGTGTGCTGTGGTAGCCAGCTTTTATTCCTTTATTTGGCCACACCCA	3840
Db	3656	CCAGAAATGGTGTGCTGTGGTAGCCAGCTTTTATTCCTTTATTTGGCCACACCCA	3715
Qy	3841	CATCCCTGCTGATTTGGCCATTTTACAGAATGCTGATTTGGTCCATTTTATAGCGTGTGAT	3900
Db	3716	CATCCCTGCTGATTTGGCCATTTTACAGAATGCTGATTTGGTCCATTTTATAGCGTGTGAT	3775
Qy	3901	TGCTGCTGTTTTTACAGAGTGTGATTTGGTGCATTTTACAAATCTTTTACGTAGACACAGAT	3960
Db	3776	TGCTGCTGTTTTTACAGAGTGTGATTTGGTGCATTTTACAAATCTTTTACGTAGACACAGAT	3835
Qy	3961	GCTGATTTGGTCCCTTTTATAATCTTTTACGTAGACACAAAAGTCTTACAAAGTCCCCACCCA	4020
Db	3836	GCTGATTTGGTCCCTTTTATAATCTTTTACGTAGACACAAAAGTCTTACAAAGTCCCCACCCA	3895
Qy	4021	ACCCAGAGCTCCGCTGGCTTCACCTCTCGCTAAGGAAATTTGAGGTTTCAACAAGTTTCAA	4080
Db	3896	ACCCAGAGCTCCGCTGGCTTCACCTCTCGCTAAGGAAATTTGAGGTTTCAACAAGTTTCAA	3955
Qy	4081	AGTCTTAAACTACAGATTTCTCATTTCTGCAACTGGATTTCCCACTCATGTGTTTGAATC	4140
Db	3956	AGTCTTAAACTACAGATTTCTCATTTCTGCAACTGGATTTCCCACTCATGTGTTTGAATC	4015
Qy	4141	CCAGGCTCTAAGACTTAACCTTGGCAATTTCTGTGACTTTATGTTCTCTGCAATTTTACAAAG	4200
Db	4016	CCAGGCTCTAAGACTTAACCTTGGCAATTTCTGTGACTTTATGTTCTCTGCAATTTTACAAAG	4075
Qy	4201	CTACTATCTGTACATCTCTGCTGTTAACTTTCAGACTAAACTTCTTTTGGATTTCAATG	4260
Db	4076	CTACTATCTGTACATCTCTGCTGTTAACTTTCAGACTAAACTTCTTTTGGATTTCAATG	4135
Qy	4261	ACCACACTTTTTTGGTTGAGGTTTGTCTATCGGTTTATTTGTTACTGTTTAAATAGAGCT	4320
Db	4136	ACCACACTTTTTTGGTTGAGGTTTGTCTATCGGTTTATTTGTTACTGTTTAAATAGAGCT	4195
Qy	4321	TCCTTCCAGAAATTTGAGTAGATGGAAGAGGAGTAGACATTTCTTAAAAATGTACCATGC	4380
Db	4196	TCCTTCCAGAAATTTGAGTAGATGGAAGAGGAGTAGACATTTCTTAAAAATGTACCATGC	4255
Qy	4381	CTTTTCAAGTCAACAGCATCCCTATCACATGCTGTCAAGGGTGGCTCAGAATAGTAGAG	4440
Db	4256	CTTTTCAAGTCAACAGCATCCCTATCACATGCTGTCAAGGGTGGCTCAGAATAGTAGAG	4315
Qy	4441	TTAAGAAATTTAAAGTAAATTTGGTGAAGCGATGAAAGCTTCATCTTAAAGCTTATATTAC	4500
Db	4316	TTAAGAAATTTAAAGTAAATTTGGTGAAGCGATGAAAGCTTCATCTTAAAGCTTATATTAC	4375
Qy	4501	ATCAACTGAATGTAAATAATTTGGAACATTTTCCAGGCATCCCTGTTATTATTATTGCT	4560
Db	4376	ATCAACTGAATGTAAATAATTTGGAACATTTTCCAGGCATCCCTGTTATTATTATTGCT	4435
Qy	4561	CTCTTTCTCTGCTTGCCTACTTCAAAAGTCATATGGCATGTGTGACTGAACACTGCTCTGCC	4620
Db	4436	CTCTTTCTCTGCTTGCCTACTTCAAAAGTCATATGGCATGTGTGACTGAACACTGCTCTGCC	4495
Qy	4621	AAAGAGTTTGTCAATAAAGATTCCTTTCTTTGTAACATTTCTACCTTGGGGCTTCAFTT	4680
Db	4496	AAAGAGTTTGTCAATAAAGATTCCTTTCTTTGTAACATTTCTACCTTGGGGCTTCAFTT	4555
Qy	4681	ATAATCAAAAGGAGTACTGTAAACCTGTCAAAAAAAGCTACCTGTGACAAATATATTATGT	4740
Db	4556	ATAATCAAAAGGAGTACTGTAAACCTGTCAAAAAAAGCTACCTGTGACAAATATATTATGT	4615
Qy	4741	GATGGTTACCTGTCAGTAAAGTGGCAATAAATAAATAAATAATACAGAAATGAAACCG	4800
Db	4616	GATGGTTACCTGTCAGTAAAGTGGCAATAAATAAATAAATAATACAGAAATGAAACCG	4675
Qy	4801	AGCAAACTGTTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTTATACAGTAA	4860
Db	4676	AGCAAACTGTTCAGAGAAATGGTCAGAAATTCACACTCTGAAGAACACGGCTTATACAGTAA	4735
Qy	4861	TAATCATATAATAGCCACTCAATCCAAAACATCCTGGCGACTTGTACATATATAA	4920
Db	4736	TAATCATATAATAGCCACTCAATCCAAAACATCCTGGCGACTTGTACATATATAA	4795
Qy	4921	TCAGTGGAGATGTGATTTGAAGCACAAGCTTTAACTGTAATGCTAGAGAGCTTAAATTTC	4980
Db	4796	TCAGTGGAGATGTGATTTGAAGCACAAGCTTTAACTGTAATGCTAGAGAGCTTAAATTTC	4855
Qy	4981	ATTTTTTGAAGAAATTTTACTTATTTTAAATGCTATCCCTGACCACTTCTGAACTTTTACTT	5040
Db	4856	ATTTTTTGAAGAAATTTTACTTATTTTAAATGCTATCCCTGACCACTTCTGAACTTTTACTT	4915
Qy	5041	GAAGATTTATTTTTTTTTTTTTTAAATACACTGTTTATTAGATTAGTATTCTGGCTTTGTT	5100
Db	4916	GAAGATTTATTTTTTTTTTTTTTAAATACACTGTTTATTAGATTAGTATTCTGGCTTTGTT	4975



repeat_region	9202..9263	/rpt_family="AT_rich"	Qy	483	GAACATAAAACAGCTCCTGGAGAGGACTATGACATCATCATGAGTGGAGTCCAGGGA	542
repeat_region	9636..9707	/rpt_family="L1"	Db	132799	GAACATAAAACAGCTCCTGGAGAGGACTATGACATCATCATGAGTGGAGTCCAGGGA	132858
repeat_region	9708..9869	/rpt_family="Alu"	Qy	543	CAGCGGACCCCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAAT	602
repeat_region	9870..10375	/rpt_family="L1"	Db	132859	CAGCGGACCCCTTTGGAAAAGGACTAGAAAGTGTGAAATCTATTAGTCTTCGATATGAAT	132918
repeat_region	10375..10748	/rpt_family="L1"	Qy	603	TCTCTGTCTGTATAAGCATTTTCATATTTACAGACACAGCGCTTCTCTAGGCGACGA	662
repeat_region	11052..11336	/rpt_family="Alu"	Db	132919	TCTCTGTCTGTATAAGCATTTTCATATTTACAGACACAGCGCTTCTCTAGGCGACGA	132978
repeat_region	11733..11939	/rpt_family="L2"	Qy	663	AAAAGTGGCAACAGCGGAGGAGGAAAAGAGATCATGAGGCATTTTCAGAGTGCACGT	722
repeat_region	12053..12183	/rpt_family="L2"	Db	132979	AAAAGTGGCAACAGCGGAGGAGGAAAAGAGATCATGAGGCATTTTCAGAGTGCACGT	133038
repeat_region	12220..12770	/rpt_family="ERV1"	Qy	723	CTTTTCATATATTTCTCAATGCCGTATGTTTGGTGTATTTTGGCCAGCATAACAATCT	782
repeat_region	12778..12994	/rpt_family="L2"	Db	133039	CTTTTCATATATTTCTCAATGCCGTATGTTTGGTGTATTTTGGCCAGCATAACAATCT	133098
repeat_region	13029..13457	/rpt_family="ERV1"	Qy	783	GCTCAAGAAAAAATCTGGAGAAAAACAAGGTGCCTTTCGCCAATGTTATGTTCTTTT	842
repeat_region	13461..13784	/rpt_family="L2"	Db	133099	GCTCAAGAAAAAATCTGGAGAAAAACAAGGTGCCTTTCGCCAATGTTATGTTCTTTT	133158
repeat_region	14947..14994	/rpt_family="(CA)n"	Qy	843	TGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGTCAATTCATTAC	902
repeat_region	16734..16769	/rpt_family="(CA)n"	Db	133159	TGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGTCAATTCATTAC	133218
repeat_region	17262..17302	/rpt_family="L1"	Qy	903	GTTGTGTGCAAAATATGATTTAAAGATACAACTTTGACAGAGCATGCTTTCCTAAGGGT	962
repeat_region	17696..18608	/rpt_family="ERVK"	Db	133219	GTTGTGTGCAAAATATGATTTAAAGATACAACTTTGACAGAGCATGCTTTCCTAAGGGT	133278
repeat_region	18651..18683	/rpt_family="(TG)n"	Qy	963	AGGCACGTGGAGGACTAAGGTAAGCATTTCTCAAGATCAGTTAAATCAAGAAAGGTCT	1022
repeat_region	20152..20271	/rpt_family="MIR"	Db	133279	AGGCACGTGGAGGACTAAGGTAAGCATTTCTCAAGATCAGTTAAATCAAGAAAGGTCT	133338
repeat_region	20443..20836	/rpt_family="L2"	Qy	1023	CTTTGCAATTCGAAATGCCCTTGTGCAAAATATTGTTATATTGATTAATTTTACACTTA	1082
repeat_region	21214..21271	/rpt_family="T-rich"	Db	133339	CTTTGCAATTCGAAATGCCCTTGTGCAAAATATTGTTATATTGATTAATTTTACACTTA	133398
repeat_region	21731..21843	/rpt_family="MIR"	Qy	1083	ATGGAACAACTTTTAACTTACAGATGAACAAACCCACAAAAGCAAAAATCAAAAGCCC	1142
repeat_region	24429..24451	/rpt_family="(TG)n"	Db	133399	ATGGAACAACTTTTAACTTACAGATGAACAAACCCACAAAAGCAAAAATCAAAAGCCC	133458
repeat_region	24512..24551	/rpt_family="AT_rich"	Qy	1143	TACCTATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTGCTTTGGGC	1202
repeat_region	24558..24712	/rpt_family="MIR"	Db	133459	TACCTATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTGCTTTGGGC	133518
repeat_region	24838..25477	/rpt_family="MERL_type"	Qy	1203	ATAAATGATAATGGAATATTCCAGGTATTGTTTAAATAGAGGCCCATCTACAAATCT	1262
repeat_region	26104..26655	/rpt_family="ERV1"	Db	133519	ATAAATGATAATGGAATATTCCAGGTATTGTTTAAATAGAGGCCCATCTACAAATCT	133578
repeat_region	26716..27433	/rpt_family="L1"	Qy	1263	TAGCAATACTTTGGATAAATCTTAAATTCAGCTGGACATGCTCTAAATGTTTTTATATA	1322
repeat_region	27659..27707	/rpt_family="AT_rich"	Db	133579	TAGCAATACTTTGGATAAATCTTAAATTCAGCTGGACATGCTCTAAATGTTTTTATATA	133638
repeat_region	28264..28456	/rpt_family="L1"	Qy	1323	CATCTTTGCTAGATAATTTCAAAATTTTAAAGTATGTAAGTAAATAGCTGTGCTGATC	1382
repeat_region	28765..28855	/rpt_family="MERL_type"	Db	133639	CATCTTTGCTAGATAATTTCAAAATTTTAAAGTATGTAAGTAAATAGCTGTGCTGATC	133698
repeat_region	28988..29238	/rpt_family="L1"	Qy	1383	AATTCAAAAACATTTACTTTCTCTAAATTTTAAAGTATGTAAGTAAATAGCTGTGCTGAT	1442
misc_feature	29346..29820	/note="similar to EST BF589745 (NID:gil1682069)"	Db	133699	AATTCAAAAACATTTACTTTCTCTAAATTTTAAAGTATGTAAGTAAATAGCTGTGCTGAT	133758
repeat_region	31395..31415		Qy	1443	ATCTACACATACAAATATAGATTTGTTTTTCATTAATAATGCTTCATCTTAAACAGAAATGT	1502
Query Match	91.3%	Score 4763.6; DB 9; Length 168595;	Db	133759	ATCTACACATACAAATATAGATTTGTTTTTCATTAATAATGCTTCATCTTAAACAGAAATGT	133818
Best Local Similarity	99.6%	pred. No. 0;	Qy	1503	CTTTGTGATTTGTTTTAGAAAACTGAGAGTTTTTAATTCATTAATTTGATCAAAAAATTT	1562
Matches 4775; Conservative	0; Mismatches	19; Indels	0; Gaps	0;	CTTTGTGATTTGTTTTAGAAAACTGAGAGTTTTTAATTCATTAATTTGATCAAAAAATTT	133878

QY	1563	GTGGGCAACATCCAGCATTAATGTTATGTTGTTTATGCTACATAAGGAGTCTTAAG	1622		134959	AGGGCAAGGAGTAGAATTAACCTCTTTAATACCTGTGCACAGTAGGCTTTTCAGA	135018	
Db	133879	GTGGGAAACAATCCAGCATTAATGTTATGTTGTTTATGCTACATAAGGAGTCTTAAG	133938		QY	2703	AACTCTACAACCCCTACATAAACTGGATAGTTAGTAAAGACACTCCCAGGAAGGCGGTT	2762
QY	1623	CTTGTCGCTTGAAGTCTTTTGTACTTAGTCCCATGTTTAAAAATPACTACTTTATPATCTA	1682		Db	135019	AACTCTACAACCCCTACATAAACTGGATAGTTAGTAAAGACACTCCCAGGAAGGCGGTT	135078
Db	133939	CTTGTCGCTTGAAGTCTTTTGTACTTAGTCCCATGTTTAAAAATPACTACTTTATPATCTA	133998		QY	2763	ATGTTTTCCAGTTTGAATCAGAAGAAATAGAGCTATAGCAATCTTCATCTCTATAGTAACAT	2822
QY	1683	AAGCATTTATGTTTTTCAATTCATTTACATGATGCTAATATATGCAATTTATAACAATA	1742		Db	135079	ATGTTTTCCAGTTTGAATCAGAAGAAATAGAGCTATAGCAATCTTCATCTCTATAGTAACAT	135138
Db	133999	AAGCATTTATGTTTTTCAATTCATTTACATGATGCTAATATATGCAATTTATAACAATA	134058		QY	2823	TAAAGAGCCTGGTTTTATATTAGCAGTCATTAAGATTTAAAAATTTACATCTTCGCCGTT	2882
QY	1743	TTAAAGATTTCGAATAGAAATATGTAATGTTTACATACATAGAAATGAAAAGTTTCATT	1802		Db	135139	TAAAGAGCCTGGTTTTATATTAGCAGTCATTAAGATTTAAAAATTTACATCTTCGCCGTT	135198
Db	134059	TTAAAGATTTCGAATAGAAATATGTAATGTTTACATACATAGAAATGAAAAGTTTCATT	134118		QY	2883	CTTCTTACTCACAGATTTTCGAGAGGTAATGTAATGATCCAGGAGGTGAGAAATCACTGCC	2942
QY	1803	TCGTAAGCAAGATGCTGGGTGAAGAGTGTCTTTGATTTGAAGATCACTAGATTTAGTAG	1862		Db	135199	CTTCTTACTCACAGATTTTCGAGAGGTAATGTAATGATCCAGGAGGTGAGAAATCACTGCC	135258
Db	134119	TCGTAAGCAAGATGCTGGGTGAAGAGTGTCTTTGATTTGAAGATCACTAGATTTAGTAG	134178		QY	2943	TTTTATAATGCGATTTAAATTCGATGAACAAAGTTTTCCAAACAAATTAACAGTAATAAAAGA	3002
QY	1863	AGGGCAAGACTTCTAGTCCCTAACTACCTTAAATAGCCATGTGCTACGTCGTAAGTCAG	1922		Db	135259	TTTTATAATGCGATTTAAATTCGATGAACAAAGTTTTCCAAACAAATTAACAGTAATAAAAGA	135318
Db	134179	AGGGCAAGACTTCTAGTCCCTAACTACCTTAAATAGCCATGTGCTACGTCGTAAGTCAG	134238		QY	3003	AACATGTATTAGCAGTTAATAAGCCAGGGGCTGGACGACGCTGTGTATCATGCTTTCAATC	3062
QY	1923	TGAACCCATCTCATCTCCTCATACTTTTTCATCTCTAAAATGAGGCTATAAATTTAAGC	1982		Db	135319	AACATGTATTAGCAGTTAATAAGCCAGGGGCTGTGTATCATGCTTTCAATC	135378
Db	134239	TGAACCCATCTCATCTCCTCATACTTTTTCATCTCTAAAATGAGGCTATAAATTTAAGC	134298		QY	3063	CATGAACCTGGTAAACTGGTACTAGTATCTCTATTGGACATGTGAGGAACCAAAATGGAGT	3122
QY	1983	TCTTCATTTTTTTTTTTTTTTTGGATAGATAGATTTTGTCTCTGTCACCCAGGTTGGAGTGCA	2042		Db	135379	CATGAACCTGGTAAACTGGTACTAGTATCTCTATTGGACATGTGAGGAACCAAAATGGAGT	135438
Db	134299	TCTTCATTTTTTTTTTTTTTTTGGATAGATAGATTTTGTCTCTGTCACCCAGGTTGGAGTGCA	134358		QY	3123	TGATAAACAGTAGAGTTAAAAATTTACTCTTCATATATTATATTGCTCTCAATCTCACAGAC	3182
QY	2043	ATGGACCATCTCAGCTCAGTCCACCTCTGCTTCCCTCGGTTCAAGTGATCTCCTGCT	2102		Db	135439	TGATAAACAGTAGAGTTAAAAATTTACTCTTCATATATTATATTGCTCTCAATCTCACAGAC	135498
Db	134359	ATGGACCATCTCAGCTCAGTCCACCTCTGCTTCCCTCGGTTCAAGTGATCTCCTGCT	134418		QY	3183	ATCTCTGCTACCAAAGCTATCATATCTAGATATCGCGCATAGGATGACCTTGGGGCAC	3242
QY	2103	TCAGCTCCCAAGTAGCGGGATTACAGGTGCGCGCCACCAACATCTGCTCAATTTTTTGT	2162		Db	135499	ATCTCTGCTACCAAAGCTATCATATCTAGATATCGCGCATAGGATGACCTTGGGGCAC	135558
Db	134419	TCAGCTCCCAAGTAGCGGGATTACAGGTGCGCGCCACCAACATCTGCTCAATTTTTTGT	134478		QY	3243	ACTAGAAATCTTTGAGAGAAATCTGGCAGAGAAACAAATATTTATTCCTACAAATAAAC	3302
QY	2163	ATTTTCACCATGTTGGCCAGGCTGTCGCAACCCCTACCTCAGGTGATCCCTCGCCTCG	2222		Db	135559	ACTAGAAATCTTTGAGAGAAATCTGGCAGAGAAACAAATATTTATTCCTACAAATAAAC	135618
Db	134479	ATTTTCACCATGTTGGCCAGGCTGTCGCAACCCCTACCTCAGGTGATCCCTCGCCTCG	134538		QY	3303	CCAGCATTTTACAGGTTTTTATTTTAACTATGAAGTATTGTTATCTGTATCTTTTCATATA	3362
QY	2223	GCTCTCAAAGTCTGGGATTACAGGTGTGAGCCACACGCCGCCCAATATCAGTTTT	2282		Db	135619	CCAGCATTTTACAGGTTTTTATTTTAACTATGAAGTATTGTTATCTGTATCTTTTCATATA	135678
Db	134539	GCTCTCAAAGTCTGGGATTACAGGTGTGAGCCACACGCCGCCCAATATCAGTTTT	134598		QY	3363	AGTGTGCCCGGAATTTATTTCTCTGTGGGTTCTTGGTCTCGCTGACTCCCAAGAAATGAA	3422
QY	2283	TCTTTTTTAAACAAGGCTAACACAATCAAAATATCTAGCTAGGGGAGAAAAAAAATAA	2342		Db	135679	AGTGTGCCCGGAATTTATTTCTCTGTGGGTTCTTGGTCTCGCTGACTCCCAAGAAATGAA	135738
Db	134599	TCTTTTTTAAACAAGGCTAACACAATCAAAATATCTAGCTAGGGGAGAAAAAAAATAA	134658		QY	3423	ACCCGAGACCTTGAGGTGAGTGTACACAGTTCTTAAAGATGGTGTTCAGAGATTGTTTC	3482
QY	2343	GGCACTGTTTATGTATACAGGCTCTTGTTCGAATCACTGGGCAGACAAATAACAGTAG	2402		Db	135739	ACCCGAGACCTTGAGGTGAGTGTACACAGTTCTTAAAGATGGTGTTCAGAGATTGTTTC	135798
Db	134659	GGCACTGTTTATGTATACAGGCTCTTGTTCGAATCACTGGGCAGACAAATAACAGTAG	134718		QY	3483	CTTCAGATGTTTCAGATGTGTCGGAGTTTCTCCCTATGTTGGTGTGAGTTCGTCCTCGCTGA	3542
QY	2403	AATCAATCCTTTTCATATATCCTCTTGCAGAAATACATAAAATCCCAAAATGGCTATCT	2462		Db	135799	CTTCAGATGTTTCAGATGTGTCGGAGTTTCTCCCTATGTTGGTGTGAGTTCGTCCTCGCTGA	135858
Db	134719	AATCAATCCTTTTCATATATCCTCTTGCAGAAATACATAAAATCCCAAAATGGCTATCT	134778		QY	3543	CTTCACCAATGAAGCCGAGACCTTTGAGTGAGTGTGTGACAGTTCCTTAAAGGACGTGC	3602
QY	2463	TCCTTTTTATGATATTTGGAGAAATGTAGCTAAGTGAAGTACAGATATTTTGGTGTATA	2522		Db	135859	CTTCACCAATGAAGCCGAGACCTTTGAGTGAGTGTGTGACAGTTCCTTAAAGGACGTGC	135918
Db	134779	TCCTTTTTATGATATTTGGAGAAATGTAGCTAAGTGAAGTACAGATATTTTGGTGTATA	134838		QY	3603	GTCCAGAGTTGTTTGTCTCCCTCCCGTAGTTCGCTCGCTGATGTCAGGAATGAAGC	3662
QY	2523	GACCACAAGGACTGTGTTGATGTTTGGCATATAAAATATACCTTAGTTTTTACTTTT	2582		Db	135919	GTCCAGAGTTGTTTGTCTCCCTCCCGTAGTTCGCTCGCTGATGTCAGGAATGAAGC	135978
Db	134839	GACCACAAGGACTGTGTTGATGTTTGGCATATAAAATATACCTTAGTTTTTACTTTT	134898		QY	3663	TGCAGACCTCGCGGTAAGTGTACAGCTCATATAAGGTAGTGCAACCCCAACACGTGAGC	3722
QY	2583	GTATGTTACATGTTAGATTTAGAGTATCAAAATTAGTAGGGAGGATTATTAACAAAGAAC	2642		Db	135979	TGCAGACCTCGCGGTAAGTGTACAGCTCATATAAGGTAGTGCAACCCCAACACGTGAGC	136038
Db	134899	GTATGTTACATGTTAGATTTAGAGTATGAAATTTAGTAGGGAGGATTATTAACAAAGAAC	134958		QY	3723	AGTAGCAAGATTTATTTATGAAGAGCAAAAGCTTCCCCCACCATTAGAAACGGACC	3782
QY	2643	AGGGCAAGGAGGTAGAAATTAACCTCTCTTAATACCTGTGCACAAGTAGGCTTTTCAGA	2702					



QY 723 CTTTTCATATATTTCTCAATGCCGTATGTTGGTGTATTTTGGCCCAAGCATACAATCT 782  
Db 27988 CTTTTCATATATTTCTCAATGCCGTATGTTGGTGTATTTTGGCCCAAGCATACAATCT 28047  
QY 783 GCTCAAGAAAAAAATCTCGAGAAAAACAAGGTGCCTTTGCCAATGTTATGTTTCTTTT 842  
Db 28048 GCTCAAGAAAAAAATCTCGAGAAAAACAAGGTGCCTTTGCCAATGTTATGTTTCTTTT 28107  
QY 843 TGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCATTCATTTAC 902  
Db 28108 TGACAAGCCCTGAGATTTCTGAGGGGAATTCACATAAATGGGATCAGGTCATTCATTTAC 28167  
QY 903 GTTGTGTGCAAAATATGATTTAAAGATACAACCTTTGACAGAGAGCATGCTTCCFAAGGT 962  
Db 28168 GTTGTGTGCAAAATATGATTTAAAGATACAACCTTTGACAGAGAGCATGCTTCCFAAGGT 28227  
QY 963 AGGCACGTGGAGGACTAAGGGTAAAGCATTTCTCAAGATCAGTTAAATCAAGAAAGGTGCT 1022  
Db 28228 AGGCACGTGGAGGACTAAGGGTAAAGCATTTCTCAAGATCAGTTAAATCAAGAAAGGTGCT 28287  
QY 1023 CTTTGCATTTCTGAATGCCCTGTGTGCAATATTTGGTTATATGATTAAATTTACACTTA 1082  
Db 28288 CTTTGCATTTCTGAATGCCCTGTGTGCAATATTTGGTTATATGATTAAATTTACACTTA 28347  
QY 1083 ATGGAACAACCTTTAACTTACAGATGAACAAACCCACAAGCAAAAAATCAAAAGCCC 1142  
Db 28348 ATGGAACAACCTTTAACTTACAGATGAACAAACCCACAAGCAAAAAATCAAAAGCCC 28407  
QY 1143 TACATATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTTTGGGC 1202  
Db 28408 TACATATGATTTTCATATTTCTGTGTAACTGGATTAAGGATTCCTGCTTTGGGC 28467  
QY 1203 ATAAATGATATGGAATTTTCCAGGTATTTGTTTAAATGAGGCCCATCTACAAATCT 1262  
Db 28468 ATAAATGATATGGAATTTTCCAGGTATTTGTTTAAATGAGGCCCATCTACAAATCT 28527  
QY 1263 TAGCAATACCTTTGATTAATCTTAAATTCAGCTGGACATCTCTAATTTGTTTTTATATA 1322  
Db 28528 TAGCAATACCTTTGATTAATCTTAAATTCAGCTGGACATCTCTAATTTGTTTTTATATA 28587  
QY 1323 CATCTTGTCTAGAAATTTCAATTTTAAATATGTGAATTTAGTTAATTAGCTGTCTGATC 1382  
Db 28588 CATCTTGTCTAGAAATTTCAATTTTAAATATGTGAATTTAGTTAATTAGCTGTCTGATC 28647  
QY 1383 AATTCAAAACATTTACTTTCTTAAATTTAGACTATGAAGTCAATAATTTCAACAATAT 1442  
Db 28648 AATTCAAAACATTTACTTTCTTAAATTTAGACTATGAAGTCAATAATTTCAACAATAT 28707  
QY 1443 ATCTACACATACAATTTATAGATTTGTTTTCATTAATGTTCTTCATCTTAACAGAAATGT 1502  
Db 28708 ATCTACACATACAATTTATAGATTTGTTTTCATTAATGTTCTTCATCTTAACAGAAATGT 28767  
QY 1503 CTTTGTGATTTGTTTTAGAAAACCTGAGAGTTTTTAATTCATAATTTACTTGATCAAAAAAT 1562  
Db 28768 CTTTGTGATTTGTTTTAGAAAACCTGAGAGTTTTTAATTCATAATTTACTTGATCAAAAAAT 28827  
QY 1563 GTGGGAACAATCCAGCATTAATTTGATGATGTTGTTTTATGTCACATAAGAGTCTTAAAG 1622  
Db 28828 GTGGGAACAATCCAGCATTAATTTGATGATGTTGTTTTATGTCACATAAGAGTCTTAAAG 28887  
QY 1623 CTTGGTGCCTTGAAGTCTTTTGTACTTTAGTCCCATGTTTAAATTTACTTTTATATCTA 1682  
Db 28888 CTTGGTGCCTTGAAGTCTTTTGTACTTTAGTCCCATGTTTAAATTTACTTTTATATCTA 28947  
QY 1683 AAGCATTTATCTTTTCAATTCATTTACATGATGCTAAATTTATGGCAATTAACAATA 1742  
Db 28948 AAGCATTTATCTTTTCAATTCATTTACATGATGCTAAATTTATGGCAATTAACAATA 29007  
QY 1743 TTAAGATTTTGGAAATAGAATATGTAATTTGTTTCACATACATAGAAATGAAAAGTTCATT 1802  
Db 29008 TTAAGATTTTGGAAATAGAATATGTAATTTGTTTCACATACATAGAAATGAAAAGTTCATT 29067

QY 1803 TCSTAAAGCAAGATGCTGGGTGAAGAGTGCTTTTGTATGTAAGATCACTAGATTAGTAG 1862  
Db 29068 TCGTAAAGCAAGATGCTGGGTGAAGAGTGCTTTTGTATGTAAGATCACTAGATTAGTAG 29127  
QY 1863 AGGCAAGACATCTTAGTCCCTTAATCTACCCCTTAATAGCCATGTGGTCACTGTAAAGTCAG 1922  
Db 29128 AGGCAAGACATCTTAGTCCCTTAATCTACCCCTTAATAGCCATGTGGTCACTGTAAAGTCAG 29187  
QY 1923 TGAACCCATCTCATCTCTCCCTCATCTATCTTTTTCATCTCTAAAATGAGGGTATAATTTAAGC 1982  
Db 29188 TGAACCCATCTCATCTCTCCCTCATCTATCTTTTTCATCTCTAAAATGAGGGTATAATTTAAGC 29247  
QY 1983 TCTTCATTTTCTTTTCTTTTGTGAGATAGAGTTTGTCTTCTGTCACCCAGGTGGAGTGCA 2042  
Db 29248 TCTTCATTTTCTTTTCTTTTGTGAGATAGAGTTTGTCTTCTGTCACCCAGGTGGAGTGCA 29307  
QY 2043 ATGCGACGATCTCAGCTCAGTCACTGCAACCTCTGTCTTCTCGGTTTCAAGTGATTCCTCGCT 2102  
Db 29308 ATGCGACGATCTCAGCTCAGTCACTGCAACCTCTGTCTTCTCGGTTTCAAGTGATTCCTCGCT 29367  
QY 2103 TCAGCCTCCCAAGTAGCCGGATTTACAGGTGCCCGCCACACACATCTGCTAATTTTGT 2162  
Db 29368 TCAGCCTCCCAAGTAGCCGGATTTACAGGTGCCCGCCACACATCTGCTAATTTTGT 29427  
QY 2163 ATTTTCACTGTTGGCCAGGCTGCTCGAACCCTTACCTCAGGTGATCCCTCGCCTCG 2222  
Db 29428 ATTTTCACTGTTGGCCAGGCTGCTCGAACCCTTACCTCAGGTGATCCCTCGCCTCG 29487  
QY 2223 GCCTCTCAAAGTGTGGGATTTACAGGTGTGAGCCACACGCCCAATATCAGTTT 2282  
Db 29488 GCCTCTCAAAGTGTGGGATTTACAGGTGTGAGCCACACGCCCAATATCAGTTT 29547  
QY 2283 TCTTTTTTAAACAAAGCTTAACACAATCAAAATTAAGTAGGGGAGAAAAATAA 2342  
Db 29548 TCTTTTTTAAACAAAGCTTAACACAATCAAAATTAAGTAGGGGAGAAAAATAA 29607  
QY 2343 GGCACTGTTATGTGAACAGGCTTGTGTCATCTACTGGGACAGACAATAAACAGTAAG 2402  
Db 29608 GGCACTGTTATGTGAACAGGCTTGTGTCATCTACTGGGACAGACAATAAACAGTAAG 29667  
QY 2403 AATCAATCCCTTTTCATATATCTCTTGTGAGAGATACATAAATCCCAAAATGGCTATCT 2462  
Db 29668 AATCAATCCCTTTTCATATATCTCTTGTGAGAGATACATAAATCCCAAAATGGCTATCT 29727  
QY 2463 TCCTTTTTATGATATTTGGAGAAATTTAGCTAAGTGACAGATATTTTCTGGGTGATATA 2522  
Db 29728 TCCTTTTTATGATATTTGGAGAAATTTAGCTAAGTGACAGATATTTTCTGGGTGATATA 29787  
QY 2523 GACCACAAAGGACGTGCTTTTCATGATGATGTTTGCATAAAATTAATACCTTAGTTTACTTT 2582  
Db 29788 GACCACAAAGGACGTGCTTTTCATGATGATGTTTGCATAAAATTAATACCTTAGTTTACTTT 29847  
QY 2583 GTATGTTTACATGTTTAGATTTAGATAAGAAATTTAGTAGGGAGGATTAATTAACAAAGAAC 2642  
Db 29848 GTATGTTTACATGTTTAGATTTAGATAAGAAATTTAGTAGGGAGGATTAATTAACAAAGAAC 29907  
QY 2643 AGGCGACAGAGGATAGAAATTAACCTCTTCTAATACCTGTGCAAGTAGGCTTTTTCAGA 2702  
Db 29908 AGGCGACAGAGGATAGAAATTAACCTCTTCTAATACCTGTGCAAGTAGGCTTTTTCAGA 29967  
QY 2703 AACTCTCAACCCCTACATAAATCTGATAGTTTGAAGAAAGACACACACACAGGCGGT 2762  
Db 29968 AACTCTCAACCCCTACATAAATCTGATAGTTTGAAGAAAGACACACACAGGCGGT 30027  
QY 2763 ATGTTTTCAGGTTTGAATCAGAGAAATAGAGCTATAGCAATCTTCTTCTATATAGTAACAT 2822  
Db 30028 ATGTTTTCAGGTTTGAATCAGAGAAATAGAGCTATAGCAATCTTCTTCTATATAGTAACAT 30087  
QY 2823 TAAAGAGCCTGGTTTATATATATAGCAGTCATTAAGATTTAAATAATTTACATCTTGGCGTT 2882  
Db 30088 TAAAGAGCCTGGTTTATATATATAGCAGTCATTAAGATTTAAATAATTTACATCTTGGCGTT 30147  
QY 2883 CTTCTTACTCACATTTTTCGAGAGGTAATGTAATGATCCAGAGGTGAGAAATCAGTCGC 2942



|||||  
Db 30148 CTTCTTACTCACAGATTTTCGAGAGGTAATGTAATGATCCAGGAGTGAGAATCACTGCC 30207  
Qy 2943 TTTTATAATCGGATTAATTTGCATGAACAAGTTTCCAAACAATAACAGTAATAAANAAGA 3002  
Db 30208 TTTTATAATCGGATTAATTTGCATGAACAAGTTTCCAAACAATAACAGTAATAAANAAGA 30267  
Qy 3003 AACATGTAATAGCACTTAATAAGCCAGGGCTGGACGACGTGTACATGCTTTCAATC 3062  
Db 30268 AACATGTAATAGCACTTAATAAGCCAGGGCTGTACGACGTGTACATGCTTTCAATC 30327  
Qy 3063 CATGAACGTGTAACCTGGTACTAGTATCTCTATTGGACATGTGAGGAACCAAAATGGAGT 3122  
Db 30328 CATGAACGTGTAACCTGGTACTAGTATCTCTATTGGACATGTGAGGAACCAAAATGGAGT 30387  
Qy 3123 TGATAAACAGTAGAGTTAAATAATTTACTCTTCATATATATTATTTGCCCTCAATCTCACAGAC 3182  
Db 30388 TGATAAACAGTAGAGTTAAATAATTTACTCTTCATATATATTATTTGCCCTCAATCTCACAGAC 30447  
Qy 3183 ATCTCTGCTACCAAAAGCTATCATATCTAGATATGGGGCATAAGGATGACCTTTGGGGCAC 3242  
Db 30448 ATCTCTGCTACCAAAAGCTATCATATCTAGATATGGGGCATAAGGATGACCTTTGGGGCAC 30507  
Qy 3243 ACTAGAAATTTCTTGAGAGAAATCTGCGAGAGAAACAAAPATTTATTCCTACAAATAAAC 3302  
Db 30508 ACTAGAAATTTCTTGAGAGAAATCTGCGAGAGAAACAAAPATTTATTCCTACAAATAAAC 30567  
Qy 3303 CCAGCATTTTACAGGTTTATTTTAACTATAGTATGTTATCTGATCTATCTTCATATA 3362  
Db 30568 CCAGCATTTTACAGGTTTATTTTAACTATAGTATGTTATCTGATCTATCTTCATATA 30627  
Qy 3363 AGTGTGCCGGAAATTTATTTCTCTGCTGGGTTCTTGGTCTCGCTGACTCCAAAGAATGAA 3422  
Db 30628 AGTGTGCCGGAAATTTATTTCTTCTGCTGGGTTCTTGGTCTCGCTGACTCCAAAGAATGAA 30687  
Qy 3423 ACCGAGACCTTGAGTGAGTGTCACAGTTCTTAAGATGGTGTTCAGAGTTGTTC 3482  
Db 30688 ACCGAGACCTTGAGTGAGTGTCACAGTTCTTAAGATGGTGTTCAGAGTTGTTC 30747  
Qy 3483 CTTTCAGATGTTACAGTTGTGTCGGAGTTCTCCCTTATGGTGAGTTCTCGCTGA 3542  
Db 30748 CTTTCAGATGTTACAGTTGTGTCGGAGTTCTCCCTTATGGTGAGTTCTCGCTGA 30807  
Qy 3543 CTTTCAACAATGAAGCCGACAGACCTTTGCAGTGAGTGTGTGACAGTTCTTAAAGCGAGTGC 3602  
Db 30808 CTTTCAACAATGAAGCCGACAGACCTTTGCAGTGAGTGTGTGACAGTTCTTAAAGCGAGTGC 30867  
Qy 3603 GTCAGAGTTGTTGTTCTCCCGGTAGGTTCTGGTCTCGCTGATGTGAGGAATGAAGC 3662  
Db 30868 GTCAGAGTTGTTGTTCTCCCGGTAGGTTCTGGTCTCGCTGATGTGAGGAATGAAGC 30927  
Qy 3663 TGCAGACCTTCGCGGTAAGTGTACAGCTCATAAAGGTAGTGCACACCCAAACAGTGAGC 3722  
Db 30928 TGCAGACCTTCGCGGTAAGTGTACAGCTCATAAAGGTAGTGCACACCCAAACAGTGAGC 30987  
Qy 3723 AGTAGCAAGATTTATTTATGAAGAGCAAAAGAAACAAAGCTTCCCCACCATAGAAACGGACC 3782  
Db 30988 AGTAGCAAGATTTATTTATGAAGAGCAAAAGAAACAAAGCTTCCCCACCATAGAAACGGACC 31047  
Qy 3783 AGAATTTGGTGTGCTGTGCTAGCCAGCTTTTATCCCTTATTTGGCCACACCCACA 3842  
Db 31048 AGAATTTGGTGTGCTGTGCTAGCCAGCTTTTATCCCTTATTTGGCCACACCCACA 31107  
Qy 3843 TCCCTGCTGATTTGGCCCATTTTACAGAAATGCTGATGTCATTTTATAGCGTCTGATTTG 3902  
Db 31108 TCCCTGCTGATTTGGCCCATTTTACAGAAATGCTGATGTCATTTTATAGCGTCTGATTTG 31167  
Qy 3903 GTGCGTTTTCACAGAGTGTGATTTGGTGTGATTTTACAAATCCTTTAGCTAGACACAGAGTGC 3962  
Db 31168 GTGCGTTTTCACAGAGTGTGATTTGGTGTGATTTTACAAATCCTTTAGCTAGACACAGAGTGC 31227  
Qy 3963 TGATTTGGTGTGCTTTTAAATCCCTTTAGCTAGACACAAAGTTCTACAAAGTCCCCACCCAC 4022  
|||||

Db 31228 TGATTTGGTGTCTTTTAAATCCTTTTAGCTAGACACAAAGTTCTACAAAGTCCCCACCCAAC 31287  
Qy 4023 CCAGAGCTTCGCTGGCTTCACCTCTCGTAAGGAATTTGAGTTTCAACAAAGTTTCAAG 4082  
Db 31288 CCAGAGCTTCGCTGGCTTCACCTCTCGTAAGGAATTTGAGTTTCAACAAAGTTTCAAG 31347  
Qy 4083 TGCTAAACCTACAGTTTCTCTCTGCAACCTGGATTTCCACTCATGTGTTGAAATCCC 4142  
Db 31348 TGCTAAACCTACAGTTTCTCTCTGCAACCTGGATTTCCACTCATGTGTTGAAATCCC 31407  
Qy 4143 AGGCTCTAAGACTTAACTTCCCATCTGTGACTTTATGTTCCCTGCAATTTACACAAAGCT 4202  
Db 31408 AGGCTCTAAGACTTAACTTCCCATCTGTGACTTTATGTTCCCTGCAATTTACACAAAGCT 31467  
Qy 4203 ACTATCTGTGCACATCTCTGGTGTAACTTCAGACTAAACTCTTTTGTGATTCACAAATGAC 4262  
Db 31468 ACTATCTGTGCACATCTCTGGTGTAACTTCAGACTAAACTCTTTTGTGATTCACAAATGAC 31527  
Qy 4263 CACACACTTTTGGTTGAGGTTTGTCTATCGGTTTATGTTACTGTTTAAATAGAGAGTTC 4322  
Db 31528 CACACACTTTTGGTTGAGGTTTGTCTATCGGTTTATGTTACTGTTTAAATAGAGAGTTC 31587  
Qy 4323 TCCAGAAATTTGAGTAGATGGAAGAGAGTAGACACATTTCTTAAAAATGTACATGCCT 4382  
Db 31588 TCCAGAAATTTGAGTAGATGGAAGAGAGTAGACACATTTCTTAAAAATGTACATGCCT 31647  
Qy 4383 TTCAGTGTCAACATCCCTTATCAGATGCTGCTCAAGGGTGGCTCAGAAATAGGTAGAGTT 4442  
Db 31648 TTCAGTGTCAACATCCCTTATCAGATGCTGCTCAAGGGTGGCTCAGAAATAGGTAGAGTT 31707  
Qy 4443 AAGAAATTTAAAGTAAATTTGGTGTAAAGCATGAAAGCTTTCATCTTAAAGCTTATATACAT 4502  
Db 31708 AAGAAATTTAAAGTAAATTTGGTGTAAAGCATGAAAGCTTTCATCTTAAAGCTTATATACAT 31767  
Qy 4503 CAACTGAAATGTAAAATTAATTTGGAACATTTTCCAGGATCCCTGTTATTTATTTGCTCT 4562  
Db 31768 CAACTGAAATGTAAAATTAATTTGGAACATTTTCCAGGATCCCTGTTATTTATTTGCTCT 31827  
Qy 4563 CTTTCTGCTGCTCCCTTACTTCAAAAGTCATATGGCATGCTGCTAGAACTGCTCGCAA 4622  
Db 31828 CTTTCTGCTGCTCCCTTACTTCAAAAGTCATATGGCATGCTGCTAGAACTGCTCGCAA 31887  
Qy 4623 AGAGTTTGTCAATATAAGATTCCTTCTTTGTAAACATTTCTACCTTTGGGCTTCATTTAT 4682  
Db 31888 AGAGTTTGTCAATATAAGATTCCTTCTTTGTAAACATTTCTACCTTTGGGCTTCATTTAT 31947  
Qy 4683 AATCAAAAGAGTACTGTAACTGTCAAAAAGTACCTGTGACAAATATATATATGTA 4742  
Db 31948 AATCAAAAGAGTACTGTAACTGTCAAAAAGTACCTGTGACAAATATATATATGTA 32007  
Qy 4743 TGGTTACCTGCAAGTGGTGGCAATAAATAAATAAATAATATATACAGAAATGAACCGAG 4802  
Db 32008 TGGTTACCTGCAAGTGGTGGCAATAAATAAATAAATAAATAATATATACAGAAATGAACCGAG 32067  
Qy 4803 CAGAACTGTCAAGAAATGTCAGAAATTCACACTCTGAAGAACACACGGCTATACAGTAATA 4862  
Db 32068 CAGAACTGTCAAGAAATGTCAGAAATTCACACTCTGAAGAACACACGGCTATACAGTAATA 32127  
Qy 4863 ATCATATAAATAGCCACTCAATCCAAAACATCAGTGGGGACTTGTACATATATAATC 4922  
Db 32128 ATCATATAAATAGCCACTCAATCCAAAACATCAGTGGGGACTTGTACATATATAATC 32187  
Qy 4923 AGTGGAGATGTGATTTGAAGCACAAGGCTTAAAGTGAATGTCTAGAGAGCTAAATGATTCAT 4982  
Db 32188 AGTGGAGATGTGATTTGAAGCACAAGGCTTAAAGTGAATGTCTAGAGAGCTAAATGATTCAT 32247  
Qy 4983 TTTTATGGAATTTTACTTTATTTTAAATGTCACTCCCTGACCACTTTTGAACCTTTTACTTGA 5042  
Db 32248 TTTTATGGAATTTTACTTTATTTTAAATGTCACTCCCTGACCACTTTTGAACCTTTTACTTGA 32307  
Qy 5043 AGATTTATTTTATTTTAAACACACTGTTTATATAGATTTAGGTATTTCTGGTCTTCTTTT 5102  
Db 32308 AGATTTATTTTATTTTAAACACACTGTTTATATAGATTTAGGTATTTCTGGTCTTCTTTT 32367





Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics -----  
Center project name: H\_RG343H22

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GT6/CHR7> or <mailto:egreen@nhgri.nih.gov>

**SOURCE INFORMATION:**  
Clone CTA-343H22 is from a release of the human BAC library C17B-HS-A. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBelobAC11  
Selection: chloram

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the right is RP11-573B16. Actual start of  
this clone is at base position 1 of CTA-343H22 actual end is at  
base position 131458 of CTA-343H22.

FEATURES	
source	base position 189490 chr1:249375000
	location/Qualifiers
	1..131458
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="CTA-343H22"
	/clone_lib="CITB-HS-A"
	32..236
repeat_region	/rpt_family="ERV1"
repeat_region	256..523
repeat_region	/rpt_family="MaLR"
repeat_region	1137..1227
repeat_region	/rpt_family="ERV1"
repeat_region	1228..1485
repeat_region	/rpt_family="Alu"
repeat_region	1486..1897
repeat_region	/rpt_family="ERV1"
repeat_region	2406..2489
repeat_region	/rpt_family="Tip100"
repeat_region	2721..2877
repeat_region	/rpt_family="L1"
repeat_region	3370..3614
repeat_region	/rpt_family="L1"
repeat_region	3619..4195
repeat_region	/rpt_family="L1"
repeat_region	4790..5239
repeat_region	/rpt_family="ERV1"
variation	4959
	/allele="n"



Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 193041)  
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and  
Haugen,E.D.  
Direct Submission  
Unpublished  
2 (bases 1 to 193041)  
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and  
Haugen,E.D.  
Direct Submission  
Submitted (14-JUN-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 193041)  
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and  
Haugen,E.D.  
Direct Submission  
Submitted (20-NOV-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Nov 20, 2001 this sequence version replaced gi:14423605.  
----- Genome Center  
Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: <http://www.genome.washington.edu>  
Contact: uwgctgs@u.washington.edu  
Drafting Center: SC  
----- Project Information  
Center project name: chr-1  
Center clone name: RP11-543B16 (sc0173)  
----- Summary Statistics  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator ET; 88% of reads  
Assembly: Dye-terminator Big Dye; 12% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 192648 bases at least Q40  
Consensus quality: 193013 bases at least Q30  
Consensus quality: 193041 bases at least Q20  
Insert size: 200871; 9.7% error; agarose-  
Insert size: 193041; sum-of-contigs  
Quality coverage: 6.5x in Q20 bases; agarose-  
Quality coverage: 6.7x in Q20 bases; sum-of-contigs  
-----  
Overlapping Sequences:  
5': Mapping in progress  
3': Mapping in progress  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were either double-stranded or sequenced with an  
alternate chemistry or covered by high quality data (i.e., Phred  
quality >= 30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by at least one plasmid subclone or more than one M13  
subclone; and the assembly was confirmed by restriction digest.  
-----  
Sequence Validation:  
This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and  
vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-800 bp)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered



2

*	1	3070:	contig of 3070 bp in length
*	3071	3170:	gap of 100 bp
*	3171	16611:	contig of 13441 bp in length
*	16612	16711:	gap of 100 bp
*	16712	26583:	contig of 9872 bp in length
*	26584	26683:	gap of 100 bp
*	26684	34934:	contig of 8251 bp in length
*	34935	35034:	gap of 100 bp
*	35035	37252:	contig of 2218 bp in length
*	37253	37352:	gap of 100 bp
*	37353	39568:	contig of 2216 bp in length
*	39569	39668:	gap of 100 bp
*	39669	59699:	contig of 17301 bp in length
*	59970	57069:	gap of 100 bp
*	57070	60245:	contig of 3176 bp in length
*	60246	60345:	gap of 100 bp





During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/ChrX>  
RP13-310B24 is from the library RPCI-13.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pBACE3.6  
IMPORTANT: This sequence is not the entire insert of clone RP13-310B24. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
The true left end of clone RP13-310B24 is at 1 in this sequence.  
The true left end of clone RP13-348B13 is at 99533 in this sequence.

## FEATURES

	Location/Qualifiers	source
	1..99632	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="X"	
	/map="q21.31-21.33"	
	/clone="RP13-310B24"	
	/clone_lib="RPCI-13.2"	
repeat_region	1..387	
	/note="TIGER1 repeat: matches 1175..1558 of consensus"	
repeat_region	388..692	
	/note="AluYa5 repeat: matches 1..305 of consensus"	
repeat_region	693..1810	
	/note="TIGER1 repeat: matches 45..1175 of consensus"	
repeat_region	1818..2361	
	/note="LIM2 repeat: matches 669..122 of consensus"	
repeat_region	3019..3339	
	/note="LIPB2 repeat: matches 5825..6154 of consensus"	
repeat_region	3395..5413	
	/note="LIM4 repeat: matches 4149..6270 of consensus"	
repeat_region	6911..7139	
	/note="L2 repeat: matches 1602..1829 of consensus"	
repeat_region	7232..7358	
	/note="MER91B repeat: matches 2..123 of consensus"	
repeat_region	7374..7711	
	/note="L2 repeat: matches 1826..2170 of consensus"	
repeat_region	8378..8873	
	/note="L2 repeat: matches 2176..2697 of consensus"	
repeat_region	9185..9476	
	/note="AluY repeat: matches 1..302 of consensus"	
repeat_region	9770..10131	
	/note="MER66A repeat: matches 2..433 of consensus"	
misc_feature	10137..10243	
	/note="match: STR: Em:L41959"	
repeat_region	11311..11561	
	/note="Tigger3(Golem) repeat: matches 2765..3027 of consensus"	
repeat_region	12387..12514	
	/note="LIMC1 repeat: matches 6175..6315 of consensus"	
repeat_region	12514..12691	
	/note="LIMC3 repeat: matches 7398..7586 of consensus"	

repeat_region	12703..12997	
	/note="AluY repeat: matches 1..292 of consensus"	
repeat_region	12999..13394	
	/note="MER31A repeat: matches 7..445 of consensus"	
repeat_region	13398..13903	
	/note="MLT2CB repeat: matches 21..491 of consensus"	
repeat_region	13936..14701	
	/note="LIMC3 repeat: matches 6625..7403 of consensus"	
repeat_region	14702..14795	
	/note="47 copies 2 mer at 64% conserved"	
repeat_region	14799..14972	
	/note="LIMC3 repeat: matches 6407..6585 of consensus"	
repeat_region	14973..15384	
	/note="MSTA repeat: matches 1..426 of consensus"	
repeat_region	15385..17011	
	/note="MSTA-internal repeat: matches 1..1651 of consensus"	
repeat_region	17015..17356	
	/note="MSTA repeat: matches 4..359 of consensus"	
repeat_region	17357..17416	
	/note="30 copies 2 mer ta 88% conserved"	
repeat_region	17474..17531	
	/note="29 copies 2 mer ac 89% conserved"	
repeat_region	17534..17585	
	/note="MSTA repeat: matches 375..426 of consensus"	
repeat_region	17609..18436	
	/note="LIMC3 repeat: matches 5521..6388 of consensus"	
repeat_region	18438..19259	
	/note="LIP4 repeat: matches 5321..6141 of consensus"	
repeat_region	19283..19303	
	/note="LIMC3 repeat: matches 5379..5398 of consensus"	
repeat_region	19304..19687	
	/note="MLT1A1 repeat: matches 1..365 of consensus"	
repeat_region	19688..20527	
	/note="LIMC3 repeat: matches 4648..5379 of consensus"	
repeat_region	20528..20827	
	/note="AluX repeat: matches 1..300 of consensus"	
repeat_region	20828..21096	
	/note="LIMC3 repeat: matches 4390..4648 of consensus"	
repeat_region	21172..21750	
	/note="LIP16 repeat: matches 5532..6157 of consensus"	
repeat_region	21752..22295	
	/note="LIP5 repeat: matches 4291..6141 of consensus"	
repeat_region	22296..24134	
	/note="LIP5 repeat: matches 4291..6141 of consensus"	
repeat_region	24125..24386	
	/note="LIP5 repeat: matches 4017..4278 of consensus"	
repeat_region	24390..24457	
	/note="MLT2D repeat: matches 486..553 of consensus"	
repeat_region	24458..25607	
	/note="LIM1 repeat: matches 4410..5403 of consensus"	
repeat_region	25731..26237	
	/note="LIM4 repeat: matches 2749..3269 of consensus"	
repeat_region	26266..26491	
	/note="LIM4 repeat: matches 11..236 of consensus"	
repeat_region	26668..27305	
	/note="LIM1 repeat: matches 1390..750 of consensus"	
repeat_region	27306..27608	
	/note="AluSg repeat: matches 1..304 of consensus"	
repeat_region	27609..28895	
	/note="LIM1 repeat: matches 750..525 of consensus"	
repeat_region	28900..28988	
	/note="L1 repeat: matches 3816..3910 of consensus"	
repeat_region	28989..29293	
	/note="AluSg repeat: matches 1..305 of consensus"	
repeat_region	29294..29448	
	/note="L1 repeat: matches 3910..4043 of consensus"	
repeat_region	29450..29484	
	/note="MER61E repeat: matches 531..565 of consensus"	
repeat_region	29505..30294	
	/note="LIMC3 repeat: matches 5176..5983 of consensus"	
repeat_region	32806..32901	
	/note="MIR repeat: matches 122..222 of consensus"	
repeat_region	33660..33985	

```
repeat_region /note="L2 repeat: matches 1737. .2088 of consensus"
34020. .34085
repeat_region /note="L2 repeat: matches 2630. .2700 of consensus"
34300. .34740
repeat_region /note="L2 repeat: matches 2281. .2710 of consensus"
34842. .35254
misc_feature /note="L16A repeat: matches 30. .445 of consensus"
35432. .35465
repeat_region /note="match: GSS: Em:AQ538246"
35466. .35840
repeat_region /note="MER39 repeat: matches 2. .385 of consensus"
complement(36285. .36420)
misc_feature /note="match: STS: Em:L77291"
38137. .38301
repeat_region /note="L1A10 repeat: matches 6157. .6318 of consensus"
38328. .39092
repeat_region /note="L1B1 repeat: matches 5396. .6158 of consensus"
39097. .39225
repeat_region /note="MER50 repeat: matches 14. .168 of consensus"
39320. .39739
repeat_region /note="MER50 repeat: matches 281. .712 of consensus"
39756. .40753
repeat_region /note="L1 repeat: matches 4445. .5419 of consensus"
40754. .41088
repeat_region /note="MER61A repeat: matches 2. .354 of consensus"
41090. .43719
repeat_region /note="MER61-internal repeat: matches 1959. .4550 of
consensus"
43720. .44403
repeat_region /note="L1R12 repeat: matches 1. .671 of consensus"
44406. .46188
repeat_region /note="MER61-internal repeat: matches 1. .1960 of
consensus"
46189. .46482
repeat_region /note="MER61A repeat: matches 13. .354 of consensus"
46490. .48685
repeat_region /note="L1 repeat: matches 2172. .4446 of consensus"

Query Match 9.5%; Score 494.6; DB 9; Length 99632;
Best Local Similarity 80.1%; Pred. No. 4.1e-79;
Matches 621; Conservative 0; Mismatches 144; Indels 10; Gaps 3;

QY 3323 TTTTAACTATGAAGTATTGTTATCTGTATCTTTCATATAGTGTCCGCGAATTTATTT 3382
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43678 TTTACAATAACGTACATAAAACCTTAATATGATTGATGTGTCCAGATTGGTTC 43737

QY 3383 CTTCTGTGGGTCTTGTGTCTCGCTGACTCCAAAGTAAAGCCGAGACCCCTTGAGGTGA 3442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43738 TTTCTGTATGGTCTTGTGTCTCAGTCAAGTAAAGCCGAGACCCCTCACTGTGA 43797

QY 3443 GTGTACAGTCTTAAAGATGTTGTTTCAGAGTTTGTCTTTCAGATGTTTCAGATGTT 3502
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43798 GTGTACAGTCTTAAAGATGTTGTTGAGTGTGTTCTTTCAGATGTTTCAGATGTT 43857

QY 3503 CGGAGTTTCTCCCTTATGTTGAGTGTGTTCTGCTGACTTCAACATGAAGCCGAG 3562
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43858 CCAGAGTTTCTCCCTTCCAGTGGTTCATGTTCTGCTGACTTCAAGTGAAGCCACAG 43917

QY 3563 ACCTTTGAGTGTGTGACAGTCTTAAAGGAGTGGTGTCCAGAGTGTGTTGTTCTCT 3622
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43918 ACCTTTGAGTGA--GTGTTACAGCTTAAAGTGGTGTGCTGCGAGTGTGTTGTTCT 43975

QY 3623 CCGGTAGTGTGCTGCTGCTGATGTCAGGATGAAGCTGCAGACCCCTCGCGTAGT 3682
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43976 CTTGTGGTGTGCTGATCTTGTGACTTTAGAGTGAAGCCACAGCCCTTCACG---GT 44031

QY 3683 GTTACAGCTATAAAGTGTGCAACCCAAACAGTGAAGTGAAGTATTTATGA 3742
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44032 GTTACAGCTATAAAGTGTGCGGACCCAAAGGTGAGCAGCAGAGATTGTTATGA 44091

QY 3743 AGAGCAAAAGAACAAAGTCTCCCAACCATAGAAAGCCGACCAAGATTGTTGCTGCTGTG 3802
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44092 AGAGCAAAAGAACAAAGTCTCCACAGCATGGAAGGGGACTGAGGGGTTGCTGCTGTG 44151
```

```
QY 3803 ----TGGTAGCCAGCTTTTATTCCCTTATTGGCCACACCCACATCCTGCTGATGGCC 3858
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44152 GCTCCGGTGGCCAGTTTTTATTCCCTTATTGGCCACCCACATCCTGCTGATGGTCC 44211

QY 3859 APTTTACAGAAATGCTGATTGGTCCATTTATATAGCGTGTGATTTGGTGGCTTTTACAGAG 3918
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44212 APTTTACAGAGTGTGATTGGTCCATTTACAGAGTCTCTGATTGGTGGCTTTTACAGAG 44271

QY 3919 TGCTGATTTGGTGCATTTTACATTCCTTTAGCTAGACAGAGTGTGATTTGGTGGCTTTAT 3978
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44272 TGCTCATTTGGTGTGTTTACATTCCTTTAGCTAGACAGAGTGTGATTTGGTGGCTTTAT 44331

QY 3979 AATCCTTTAGCTAGACACAAAAGTTCTCAAGTCCCAAGTCCCAAGTCCCGCTGG 4038
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44332 AATCCTTTAGCTAGACACAAAAGTTCTCAAGTCCCAAGTCCCAAGTCCCGCTGG 44391

QY 4039 CTTACCTCTCTGTAAGGAAATTTAGTTCACAAAGTTCAAAAGTGTCTAAACTA 4093
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44392 CTTACCTTTCAAAAGCATATACTCAGACATTAGAATTTAGAAATCCCATACAA 44446
```

```
RESULT 11
HS30G7/c
LOCUS
DEFINITION
Human DNA sequence from clone Rp1-30G7 on chromosome Xp11.3-11.4,
complete sequence.
ACCESSION
AL034402
VERSION
AL034402.9 GI:4586155
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 123637)
AUTHORS
Whitehead,S.
TITLE
Direct Submission
JOURNAL
Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Apr 16, 1999 this sequence version replaced gi:4582299.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP database can be found at <http://www.sanger.ac.uk/projects/c-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RPI-30G7 is from the library RPI-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

[illegible]



QY 3881 CCATTTTATAGCGTCTGATTGGTGGCTTTTACAGAGTGTGATTGGTCATTACAAT 3940  
 Db 116886 CCATTTTACAGAGTCTGATTGGTGGCTTTTACAGAGTGTGATTGGTCATTACAAT 116945  
 QY 3941 CTTTACAGACAGAGTGTGATTGGTGGCTTTTATATCCTTTACAGACACAAA 4000  
 Db 116946 ACTTTCCTAGACACAGAGTGTGATTGGTGGCTTTTACATCTTTTACAGACACAAA 117005  
 QY 4001 GTTCTACAAGTCCCAACCCAGAGCTCCGCTGGCTTCACCTCTC 4049  
 Db 117006 GTTCTCAAGTCCCAACCCAGAGCTCCGCTGGCTTCACCTCTC 117054

RESULT 14  
 AC104697/c  
 LOCUS AC104697 159797 bp DNA linear PRI 21-FEB-2002  
 DEFINITION Homo sapiens BAC clone RP11-454B3 from 2, complete sequence.  
 ACCESSION AC104697 AC024403  
 VERSION AC104697.2 GI:18855149  
 KEYWORDS HTG.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 159797)  
 Tilden, J.E. and Waterston, R.  
 Toward a complete human genome sequence  
 Genome Res. 8 (11), 1097-1108 (1998)  
 93063792  
 PUBMED 9847074

REFERENCE 2 (bases 1 to 159797)  
 Radonkenko, M. and Cotton, M.  
 The sequence of Homo sapiens BAC clone RP11-454B3  
 Unpublished (2001)  
 3 (bases 1 to 159797)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (19-DEC-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 4 (bases 1 to 159797)  
 Waterston, R.  
 Direct Submission  
 Submitted (21-FEB-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Feb 21, 2002 this sequence version replaced gi:17933873.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: sapiens@wustl.edu  
 ----- Summary Statistics  
 Center project name: H\_NH0454B03  
 Drafting Center: WIBR  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this

sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenko, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACE3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-337B18, 2000 bp overlap; the clone sequenced to the right is RP11-241F6, 2000 bp overlap. Actual start of this clone is at base position 167881 of RP11-337B18; actual end is at base position 2890 of RP11-241F6.

Polymorphisms have been identified between AC092632, AC096560, and AC024403.

The sequence of AC024403 has been incorporated into AC104697.

FEATURES	Location/Qualifiers
source	1. 159797
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-454B3"
	/clone_lib="RPCI-11"
	340..410
repeat_region	/rpt_family="(TA)n"
	492..514
repeat_region	/rpt_family="AT-rich"
	1713..1931
repeat_region	/rpt_family="MIR"
	2128..2272
repeat_region	/rpt_family="MER1_type"
	2873..2998
repeat_region	/rpt_family="L2"
	3102..3144
repeat_region	/rpt_family="T-rich"
	3113..3425
repeat_region	/rpt_family="Alu"
	4125..4273
repeat_region	/rpt_family="MER1_type"
	5225..5808
repeat_region	/rpt_family="L1"
	5811..7833
repeat_region	/rpt_family="L1"
	7068..7094
repeat_region	/rpt_family="AT-rich"
	8088..8382
repeat_region	/rpt_family="Alu"
	8742..8807
repeat_region	/rpt_family="AT-rich"
	8900..9206
repeat_region	/rpt_family="Alu"
	9180..9208
repeat_region	/rpt_family="AT-rich"
	9230..9332
repeat_region	/rpt_family="MIR"
	11500..11687
repeat_region	/rpt_family="Alu"
	11680..11714
repeat_region	/rpt_family="(TTTA)n"
	11688..11998
repeat_region	/rpt_family="Alu"
	11999..12083
repeat_region	/rpt_family="Alu"
	12622..13188
repeat_region	/rpt_family="L1"



